Journal of Data Science **12**(2014), 87-106

On Chen *et al.*'s Extreme Value Distribution

Saralees Nadarajah* and Sumaya Eljabri
University of Manchester

Abstract: Chen, Bunce and Jiang [In: *Proceedings of the International Conference on Computational Intelligence and Software Engineering*, pp. 1-4] claim to have proposed a new extreme value distribution. But the formulas given for the distribution do not form a valid probability distribution. Here, we correct their formulas to form a valid probability distribution. For this valid distribution, we provide a comprehensive treatment of mathematical properties, estimate parameters by the method of maximum likelihood and provide the observed information matrix. The flexibility of the distribution is illustrated using a real data set.

Key words: Extreme values, maximum likelihood, moments.

1. Introduction

The generalized extreme value (GEV) distribution is one of the most widely applied models for univariate extreme values. Its cumulative distribution function and probability density function are specified by

$$F(x) = \exp(-u),$$

and

$$f(x) = \sigma^{-1} u^{1+\xi} \exp(-u), \tag{1}$$

respectively, where $1 + \xi(x - \mu)/\sigma > 0$, $-\infty < \xi < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$ and $u = \{1 + \xi(x - \mu)/\sigma\}^{-1/\xi}$. Of the three parameters, μ is referred to as the location parameter, σ is referred to as the scale parameter, and ξ is referred to as the shape parameter. For details on the GEV distribution, its theory and applications, we refer the readers to Leadbetter *et al.* (1987), Embrechts *et al.* (1997), Castillo *et al.* (2005), and Resnick (2008).

*Corresponding author.

In recent years, several extensions of the GEV distribution have been proposed in the literature. The most recent of these is a four-parameter distribution due to Chen *et al.* (2010). Earlier generalizations include the three-parameter kappa distribution due to Mielke (1973) and the four-parameter kappa distribution due to Hosking (1994). Chen's generalization has the cumulative distribution function and the probability density function given by

$$F(x) = \left\{ 1 + \exp\left[-\frac{1}{\delta} (x-\mu)^{\alpha} \right] \right\}^{-1/\beta}, \tag{2}$$

and

$$f(x) = \alpha (\delta\beta)^{-1} (x-\mu)^{\alpha-1} \exp\left[-\frac{1}{\delta} (x-\mu)^{\alpha} \right] \left\{ 1 + \exp\left[-\frac{1}{\delta} (x-\mu)^{\alpha} \right] \right\}^{-1/\beta-1}, \tag{3}$$

respectively, for $-\infty < x < \infty$, $\alpha > 0$, $\beta > 0$, $\delta > 0$ and $-\infty < \mu < \infty$. An excellent motivation for introducing (2) and (3) is described in Chen *et al.* (2010). In spite of that, neither of (2) and (3) appear to be valid functions since $(x-\mu)^{\alpha}$ is undefined for $x < \mu$. Furthermore, the distribution given by (2) and (3) is not a generalization of the GEV distribution, so it cannot be an extreme value distribution.

The aim of this note is to provide a modification so that Chen *et al.* (2010)'s distribution becomes a valid probability distribution. The modification has the cumulative distribution function and the probability density function specified by

$$F(x) = \left(1 - 2^{-1/\beta} \right)^{-1} \left[\left\{ 1 + \exp\left[-\frac{1}{\delta} (x-\mu)^{\alpha} \right] \right\}^{-1/\beta} - 2^{-1/\beta} \right], \tag{4}$$

and

$$\begin{aligned} f(x) = {} & \alpha (\delta\beta)^{-1} \left(1 - 2^{-1/\beta} \right)^{-1} (x-\mu)^{\alpha-1} \exp\left[-\frac{1}{\delta} (x-\mu)^{\alpha} \right] \\ & \times \left\{ 1 + \exp\left[-\frac{1}{\delta} (x-\mu)^{\alpha} \right] \right\}^{-1/\beta-1}, \end{aligned} \tag{5}$$

respectively, for $\alpha > 0$, $\beta > 0$, $\delta > 0$ and $x > \mu > -\infty$. Clearly, (4) and (5) form a valid probability distribution. We shall refer to the distribution given by (4) and (5) as the Chen distribution. Although this distribution is still not an extreme value distribution, we shall see later that it can be a good competitor to the GEV distribution.

If X is a random variable with probability density function (5), we write $X \sim$ Chen$(\alpha,\ \beta,\ \delta,\ \mu)$. The Chen quantile function is obtained by inverting (4)

$$x = Q(z) = F^{-1}(z) = \mu + \left[-\delta \ln \left\{ \left[2^{-1/\beta} + \left(1 - 2^{-1/\beta} \right) z \right]^{-\beta} - 1 \right\} \right]^{1/\alpha}. \quad (6)$$

So, one can generate Chen variates from (6) by $X = Q(U)$, where U is a uniform variate on the unit interval $(0, 1)$.

In the rest of this paper, we provide a comprehensive description of the mathematical properties of (5). We examine the shape of (5) and its associated hazard rate function in Sections 2 and 3, respectively. We derive expressions for the moments in Section 4. Order statistics, their moments and L-moments are calculated in Section 5. Asymptotic distributions of the extreme values are provided in Section 6. Estimation by the method of maximum likelihood – including the observed information matrix – is presented in Section 7. A simulation study is presented in Section 8 to assess the performance of the maximum likelihood estimators. Application of the Chen distribution to a real data set is illustrated in Section 9.

The results in Section 4 involve infinite series representations. The terms of these infinite series are elementary, so the infinite series can be computed by truncation using any standard package, perhaps even pocket calculators.

2. Shape of Probability Density Function

The first derivative of $\ln\{f(x)\}$ for the Chen distribution is:

$$\frac{d \ln f(x)}{dx} = -\frac{\alpha}{\delta}(x-\mu)^{\alpha-1} + \frac{\alpha-1}{x-\mu} + \frac{\alpha}{\delta}\left(\frac{1}{\beta}+1\right) \frac{(x-\mu)^{\alpha-1}}{1+\exp\left[(x-\mu)^{\alpha}/\delta\right]}.$$

So, the modes of $f(x)$ are the roots of the equation

$$\frac{\alpha}{\delta}(x-\mu)^{\alpha} - \frac{\alpha}{\delta}\left(\frac{1}{\beta}+1\right) \frac{(x-\mu)^{\alpha}}{1+\exp\left[(x-\mu)^{\alpha}/\delta\right]} = \alpha - 1. \quad (7)$$

There may be more than one root to (7). If $x = x_0$ is a root of (7) then it corresponds to a local maximum if $d \ln f(x)/dx > 0$ for all $x < x_0$ and $d \ln f(x)/dx < 0$ for all $x > x_0$. It corresponds to a local minimum if $d \ln f(x)/dx < 0$ for all $x < x_0$ and $d \ln f(x)/dx > 0$ for all $x > x_0$. It corresponds to a point of inflexion if either $d \ln f(x)/dx > 0$ for all $x \neq x_0$ or $d \ln f(x)/dx < 0$ for all $x \neq x_0$.

Plots of the shapes of (5) for $\mu = 0$, $\delta = 1$ and selected values of (α, β) are given in Figure 1. Both unimodal and monotonically decreasing shapes appear possible. Unimodal shapes appear for large α. Monotonically decreasing shapes appear for small α.

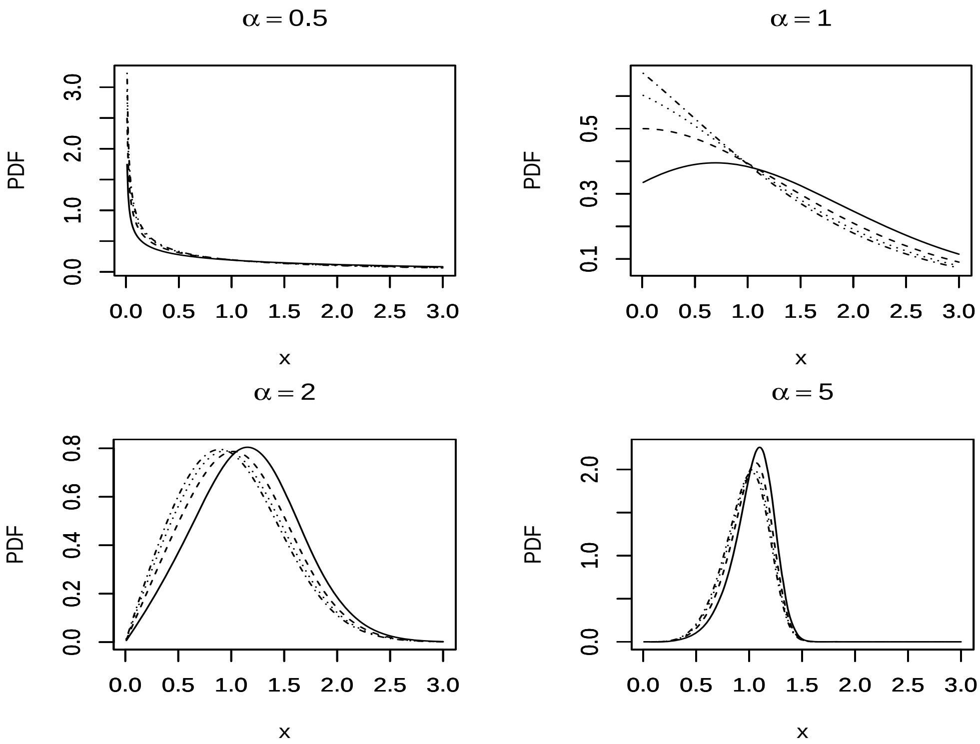

Figure 1: Plots of (5) for $\mu = 0$, $\delta = 1$, $\alpha = 0.5, 1, 2, 5$, $\beta = 0.5$ (solid curve), $\beta = 1$ (curve of dashes), $\beta = 2$ (curve of dots) and $\beta = 5$ (curve of dots and dashes

Furthermore, the asymptotes of $f(x)$ and $F(x)$ as $x \to \infty, \mu$ are given by

$$f(x) \sim \alpha(\delta\beta)^{-1} \left(1 - 2^{-1/\beta}\right)^{-1} x^{\alpha-1} \exp\left[-\frac{x^\alpha}{\delta}\right],$$

as $x \to \infty$,

$$f(x) \sim \alpha(\delta\beta)^{-1} 2^{-1/\beta-1} \left(1 - 2^{-1/\beta}\right)^{-1} (x-\mu)^{\alpha-1},$$

as $x \to \mu$,

$$1 - F(x) \sim \beta \left(1 - 2^{-1/\beta}\right)^{-1} \exp\left[-\frac{x^\alpha}{\delta}\right],$$

as $x \to \infty$, and

$$F(x) \sim (\delta\beta)^{-1} 2^{-1/\beta-1} \left(1 - 2^{-1/\beta}\right)^{-1} (x-\mu)^{\alpha},$$

as $x \to \mu$. Note that the upper tail of $f(x)$ is that of a Weibull distribution with shape parameter α and scale parameter δ. The lower tail of $f(x)$ is polynomial with power $\alpha - 1$. The upper tail of $1 - F(x)$ is that of a Weibull distribution with shape parameter α and scale parameter δ. The lower tail of $F(x)$ is polynomial with power α.

3. Shape of Hazard Rate Function

The hazard rate function defined by $h(x) = f(x)/\{1 - F(x)\}$ is an important quantity characterizing life phenomena of a system. For the Chen distribution, $h(x)$ takes the form

$$h(x) = \alpha \frac{(x-\mu)^{\alpha-1} \exp\left[-(x-\mu)^{\alpha}/\delta\right] \left\{1 + \exp\left[-(x-\mu)^{\alpha}/\delta\right]\right\}^{-1/\beta-1}}{\delta\beta \left[1 - \left\{1 + \exp\left[-(x-\mu)^{\alpha}/\delta\right]\right\}^{-1/\beta}\right]}. \tag{8}$$

The first derivative of $\ln h(x)$ is:

$$\begin{aligned}\frac{d \ln h(x)}{dx} = & -\frac{\alpha}{\delta}(x-\mu)^{\alpha-1} + \frac{\alpha-1}{x-\mu} + \frac{\alpha}{\delta}\left(\frac{1}{\beta}+1\right) \frac{(x-\mu)^{\alpha-1}}{1+\exp\left[(x-\mu)^{\alpha}/\delta\right]} \\ & + \alpha \frac{(x-\mu)^{\alpha-1} \exp\left[-(x-\mu)^{\alpha}/\delta\right] \left\{1 + \exp\left[-(x-\mu)^{\alpha}/\delta\right]\right\}^{-1/\beta-1}}{\delta\beta \left[1 - \left\{1 + \exp\left[-(x-\mu)^{\alpha}/\delta\right]\right\}^{-1/\beta}\right]}.\end{aligned}$$

So, the modes of $h(x)$ are the roots of the equation

$$\begin{aligned}& \frac{\alpha}{\delta}(x-\mu)^{\alpha} - \frac{\alpha}{\delta}\left(\frac{1}{\beta}+1\right) \frac{(x-\mu)^{\alpha}}{1+\exp\left[(x-\mu)^{\alpha}/\delta\right]} \\ & -\alpha \frac{(x-\mu)^{\alpha} \exp\left[-(x-\mu)^{\alpha}/\delta\right] \left\{1 + \exp\left[-(x-\mu)^{\alpha}/\delta\right]\right\}^{-1/\beta-1}}{\delta\beta \left[1 - \left\{1 + \exp\left[-(x-\mu)^{\alpha}/\delta\right]\right\}^{-1/\beta}\right]} = \alpha - 1.\end{aligned} \tag{9}$$

There may be more than one root to (9). If $x = x_0$ is a root of (9) then it corresponds to a local maximum if $d \ln h(x)/dx > 0$ for all $x < x_0$ and $d \ln h(x)/dx < 0$ for all $x > x_0$. It corresponds to a local minimum if $d \ln h(x)/dx < 0$ for all $x < x_0$ and $d \ln h(x)/dx > 0$ for all $x > x_0$. It corresponds to a point of inflexion if either $d \ln h(x)/dx > 0$ for all $x \neq x_0$ or $d \ln h(x)/dx < 0$ for all $x \neq x_0$.

Furthermore, the asymptotes of $h(x)$ as $x \to \infty, \mu$ are given by

$$h(x) \sim \alpha \delta^{-1} x^{\alpha-1},$$

as $x \to \infty$, and

$$h(x) \sim \alpha (\delta\beta)^{-1} 2^{-1/\beta-1} \left(1 - 2^{-1/\beta}\right)^{-1} (x-\mu)^{\alpha-1},$$

as $x \to \mu$. Note that both the upper and lower tails of $h(x)$ behave polynomially with respect to x.

Figure 2 illustrates some of the possible shapes of $h(x)$ for $\mu = 0$, $\delta = 1$ and selected values of (α, β). Both monotonically increasing, monotonically decreasing and upside down bathtub shapes appear possible. Upside down bathtub shapes appear for small values of α and β. Monotonically decreasing shapes appear for small α. Monotonically increasing shapes appear for large α.

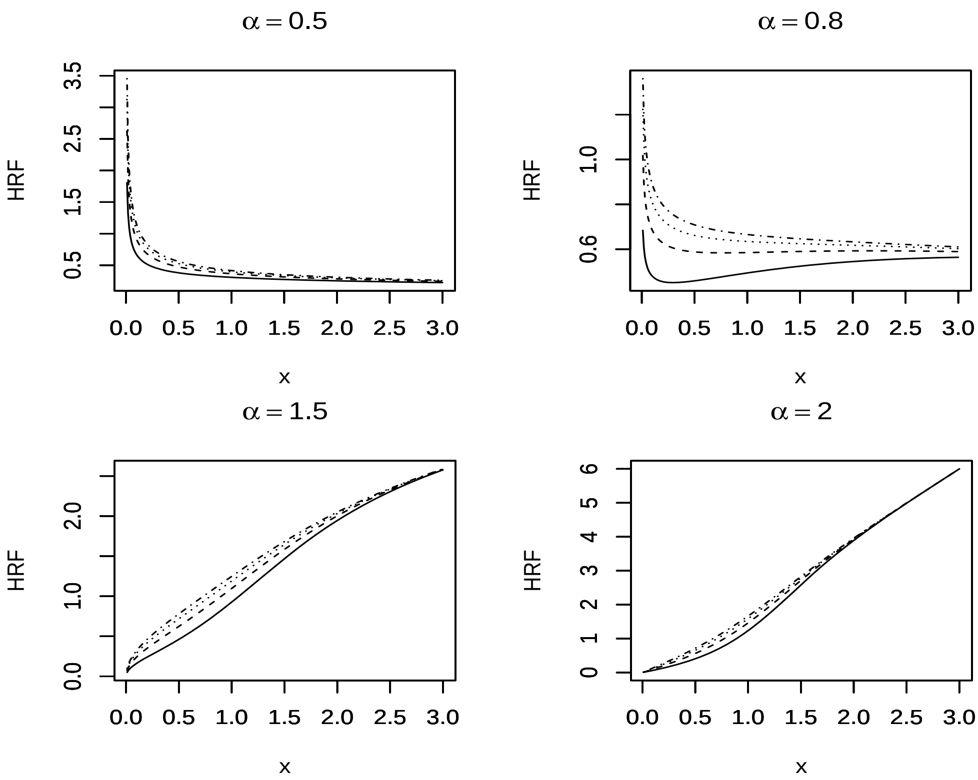

Figure 2: Plots of (8) for $\mu = 0$, $\delta = 1$, $\alpha = 0.5, 0.8, 1.5, 2$, $\beta = 0.5$ (solid curve), $\beta = 1$ (curve of dashes), $\beta = 2$ (curve of dots) and $\beta = 5$ (curve of dots and dashes)

Upside down bathtub shaped hazard rates are common in reliability and survival analysis. For example, such hazard rates can be observed in the course of a disease whose mortality reaches a peak after some finite period and then declines gradually (Silva *et al.*, 2010). For other practical examples yielding upside down bathtub hazard rates, see Singh and Misra (1994).

It is interesting to note that the Chen distribution can exhibit upside down bathtub shapes. The GEV distribution cannot exhibit upside down bathtub shaped hazard rates.

4. Moments

Let $X \sim \text{Chen}(\alpha, \beta, \delta, \mu)$. Using binomial expansion, we can write

$$
\begin{aligned}
E\left(X^n\right) &= E\left((X-\mu+\mu)^n\right) \\
&= \sum_{m=0}^{n} \binom{n}{m} \mu^{n-m} E\left((X-\mu)^m\right) \\
&= \alpha(\delta\beta)^{-1}\left(1-2^{-1/\beta}\right)^{-1} \sum_{m=0}^{n} \binom{n}{m} \mu^{n-m} \\
&\quad \times \int_{\mu}^{\infty} (x-\mu)^{m+\alpha-1} \exp\left[-\frac{1}{\delta}(x-\mu)^{\alpha}\right] \left\{1+\exp\left[-\frac{1}{\delta}(x-\mu)^{\alpha}\right]\right\}^{-1/\beta-1} dx \\
&= \alpha(\delta\beta)^{-1}\left(1-2^{-1/\beta}\right)^{-1} \sum_{m=0}^{n} \binom{n}{m} \mu^{n-m} \\
&\quad \times \int_{\mu}^{\infty} (x-\mu)^{m+\alpha-1} \exp\left[-\frac{1}{\delta}(x-\mu)^{\alpha}\right] \sum_{k=0}^{\infty} \binom{-1/\beta-1}{k} \exp\left[-\frac{k}{\delta}(x-\mu)^{\alpha}\right] dx \\
&= \alpha(\delta\beta)^{-1}\left(1-2^{-1/\beta}\right)^{-1} \sum_{m=0}^{n} \binom{n}{m} \sum_{k=0}^{\infty} \mu^{n-m} \binom{-1/\beta-1}{k} \\
&\quad \times \int_{\mu}^{\infty} (x-\mu)^{m+\alpha-1} \exp\left[-\frac{k+1}{\delta}(x-\mu)^{\alpha}\right] dx \\
&= \beta^{-1}\left(1-2^{-1/\beta}\right)^{-1} \sum_{m=0}^{n} \binom{n}{m} \sum_{k=0}^{\infty} \mu^{n-m} \binom{-1/\beta-1}{k} \delta^{m/\alpha}(k+1)^{-m/\alpha-1} \\
&\quad \times \int_{0}^{\infty} y^{m/\alpha} \exp(-y)\, dy \\
&= \beta^{-1}\left(1-2^{-1/\beta}\right)^{-1} \\
&\quad \times \sum_{m=0}^{n} \binom{n}{m} \sum_{k=0}^{\infty} \mu^{n-m} \binom{-1/\beta-1}{k} \delta^{m/\alpha}(k+1)^{-m/\alpha-1} \Gamma\left(m/\alpha+1\right) \qquad (10)
\end{aligned}
$$

for any positive integer n. The first four moments are:

$$
\begin{aligned}
E(X) = \beta^{-1}\left(1-2^{-1/\beta}\right)^{-1} \Bigg[& \mu\beta\left(1-2^{-1/\beta}\right) \\
& + \sum_{k=0}^{\infty} \binom{-1/\beta-1}{k} \delta^{1/\alpha}(k+1)^{-1/\alpha-1} \Gamma\left(1/\alpha+1\right) \Bigg], \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
E\left(X^2\right) = \beta^{-1}\left(1-2^{-1/\beta}\right)^{-1} & \left[\mu^2\beta\left(1-2^{-1/\beta}\right)\right. \\
& +2\sum_{k=0}^{\infty}\mu\binom{-1/\beta-1}{k}\delta^{1/\alpha}(k+1)^{-1/\alpha-1}\Gamma\left(1/\alpha+1\right) \\
& \left.+\sum_{k=0}^{\infty}\binom{-1/\beta-1}{k}\delta^{2/\alpha}(k+1)^{-2/\alpha-1}\Gamma\left(2/\alpha+1\right)\right], \qquad (12)
\end{aligned}
$$

$$
\begin{aligned}
E\left(X^3\right) = \beta^{-1}\left(1-2^{-1/\beta}\right)^{-1} & \left[\mu^3\beta\left(1-2^{-1/\beta}\right)\right. \\
& +3\sum_{k=0}^{\infty}\mu^2\binom{-1/\beta-1}{k}\delta^{1/\alpha}(k+1)^{-1/\alpha-1}\Gamma\left(1/\alpha+1\right) \\
& +3\sum_{k=0}^{\infty}\mu\binom{-1/\beta-1}{k}\delta^{2/\alpha}(k+1)^{-2/\alpha-1}\Gamma\left(2/\alpha+1\right) \\
& \left.+\sum_{k=0}^{\infty}\binom{-1/\beta-1}{k}\delta^{3/\alpha}(k+1)^{-3/\alpha-1}\Gamma\left(3/\alpha+1\right)\right], \qquad (13)
\end{aligned}
$$

and

$$
\begin{aligned}
E\left(X^4\right) = \beta^{-1}\left(1-2^{-1/\beta}\right)^{-1} & \left[\mu^4\beta\left(1-2^{-1/\beta}\right)\right. \\
& +4\sum_{k=0}^{\infty}\mu^3\binom{-1/\beta-1}{k}\delta^{1/\alpha}(k+1)^{-1/\alpha-1}\Gamma\left(1/\alpha+1\right) \\
& +6\sum_{k=0}^{\infty}\mu^2\binom{-1/\beta-1}{k}\delta^{2/\alpha}(k+1)^{-2/\alpha-1}\Gamma\left(2/\alpha+1\right) \\
& +4\sum_{k=0}^{\infty}\mu\binom{-1/\beta-1}{k}\delta^{3/\alpha}(k+1)^{-3/\alpha-1}\Gamma\left(3/\alpha+1\right) \\
& \left.+\sum_{k=0}^{\infty}\binom{-1/\beta-1}{k}\delta^{4/\alpha}(k+1)^{-4/\alpha-1}\Gamma\left(4/\alpha+1\right)\right]. \qquad (14)
\end{aligned}
$$

The infinite series in (10)-(14) all converge.

The expressions given by (11)-(14) can be used to compute the mean, variance, skewness and kurtosis of X. The values of these four quantities versus α are plotted in Figure 3 for $\mu = 0$, $\delta = 1$ and selected values of β. We can see that: (i) mean, variance and skewness are monotonic decreasing functions of α; (ii) kurtosis initially decreases before increasing with respect to α.

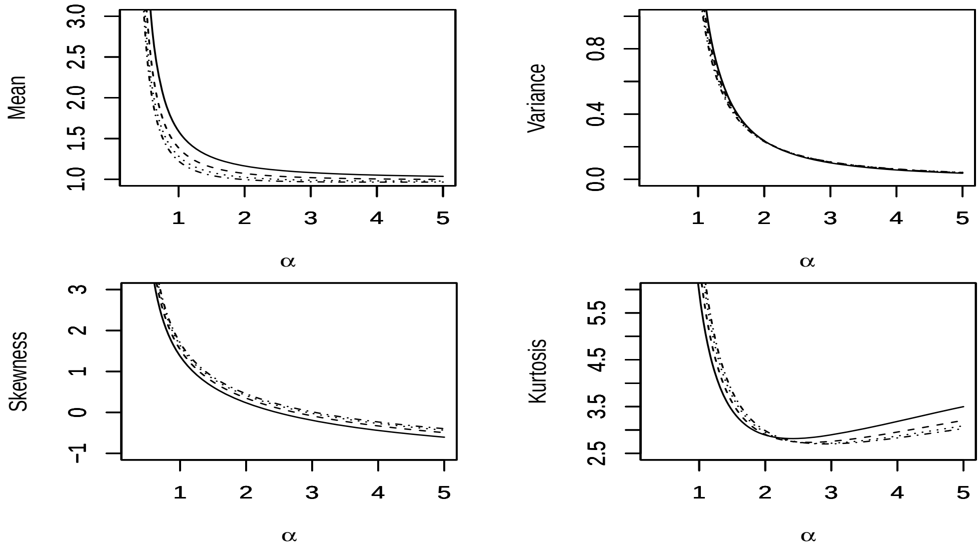

Figure 3: Mean, variance, skewness and kurtosis versus α for $\mu = 0$, $\delta = 1$, $\beta = 0.5$ (solid curve), $\beta = 1$ (curve of dashes), $\beta = 2$ (curve of dots) and $\beta = 5$ (curve of dots and dashes)

5. Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Let $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ denote the order statistics for a random sample $X_1, X_2, \cdots, X_n$ from (5). It is well known that the probability density function of the kth order statistic, say $Y = X_{k:n}$, is

$$f_Y(y) = \frac{n!}{(k-1)!(n-k)!} F^{k-1}(y) \left[1 - F(y)\right]^{n-k} f(y).$$

Substituting the expressions for $F(y)$ and $f(y)$ given by (4) and (5), respectively, we obtain

$$\begin{aligned} f_Y(y) = {} & \frac{\alpha n! \left(1 - 2^{-1/\beta}\right)^{-n}}{\delta\beta(k-1)!(n-k)!} (y-\mu)^{\alpha-1} \exp\left[-\frac{1}{\delta}(y-\mu)^{\alpha}\right] \\ & \times \left\{1 + \exp\left[-\frac{1}{\delta}(y-\mu)^{\alpha}\right]\right\}^{-1/\beta-1} \\ & \times \left[\left\{1 + \exp\left[-\frac{1}{\delta}(y-\mu)^{\alpha}\right]\right\}^{-1/\beta} - 2^{-1/\beta}\right]^{k-1} \\ & \times \left[1 - \left\{1 + \exp\left[-\frac{1}{\delta}(y-\mu)^{\alpha}\right]\right\}^{-1/\beta}\right]^{n-k} \end{aligned}$$

$$= \frac{\alpha n! \left(1 - 2^{-1/\beta}\right)^{-n}}{\delta\beta(k-1)!(n-k)!} \sum_{i=0}^{k-1} \sum_{j=0}^{n-k} \binom{k-1}{i} \binom{n-k}{j} (-1)^{k-i-1+j} 2^{(i+1-k)/\beta}$$

$$\times (y-\mu)^{\alpha-1} \exp\left[-\frac{1}{\delta}(y-\mu)^{\alpha}\right] \left\{1 + \exp\left[-\frac{1}{\delta}(y-\mu)^{\alpha}\right]\right\}^{-(i+j+1)/\beta-1}$$

$$= \frac{n! \left(1 - 2^{-1/\beta}\right)^{-n}}{(k-1)!(n-k)!}$$

$$\times \sum_{i=0}^{k-1} \sum_{j=0}^{n-k} \binom{k-1}{i} \binom{n-k}{j} \frac{2^{(i+1-k)/\beta} - 2^{-(j+k)/\beta}}{(-1)^{k-i-1+j}(i+j+1)} f_{\alpha,\beta/(i+j+1),\delta,\mu}(y),$$

where $f_{a,b,\sigma,\xi}(\cdot)$ denotes the probability density function of Chen (a, b, σ, ξ). So, the probability density function of Y is a finite linear combination of probability density functions of Chen random variables. Hence, other properties of Y can be easily derived. For instance, the cumulative distribution function of Y can be expressed as

$$F_Y(y) = \frac{n! \left(1 - 2^{-1/\beta}\right)^{-n}}{(k-1)!(n-k)!}$$

$$\times \sum_{i=0}^{k-1} \sum_{j=0}^{n-k} \binom{k-1}{i} \binom{n-k}{j} \frac{2^{(i+1-k)/\beta} - 2^{-(j+k)/\beta}}{(-1)^{k-i-1+j}(i+j+1)} F_{\alpha,\beta/(i+j+1),\delta,\mu}(y),$$

where $F_{a,b,\sigma,\xi}(\cdot)$ denotes the cumulative distribution function corresponding to $f_{a,b,\sigma,\xi}(\cdot)$. The qth moment of Y can be expressed as

$$E\left[Y^q\right] = \frac{n! \left(1 - 2^{-1/\beta}\right)^{-n}}{(k-1)!(n-k)!}$$

$$\times \sum_{i=0}^{k-1} \sum_{j=0}^{n-k} \binom{k-1}{i} \binom{n-k}{j} \frac{2^{(i+1-k)/\beta} - 2^{-(j+k)/\beta}}{(-1)^{k-i-1+j}(i+j+1)} E\left[X^q_{\alpha,\beta/(i+j+1),\delta,\mu}\right], \quad (15)$$

where $X_{a,b,\sigma,\xi} \sim$ Chen (a, b, σ, ξ).

L-moments are summary statistics for probability distributions and data samples (Hoskings, 1990). They are analogous to ordinary moments but are computed from linear functions of the ordered data values. The rth L-moment is defined by

$$\lambda_r = \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} \binom{r-1+j}{j} \beta_j,$$

where $\beta_j = E\{XF(X)^j\}$. In particular, $\lambda_1 = \beta_0$, $\lambda_2 = 2\beta_1 - \beta_0$, $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$ and $\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$. In general, $\beta_r = (r+1)^{-1} E(X_{r+1:r+1})$, so it can be computed using (15). The L-moments have several advantages over ordinary moments: for example, they apply for any distribution having finite mean; no higher-order moments need be finite.

6. Extreme Values

Suppose $X_1, \cdots, X_n$ is a random sample from (5). If $\overline{X} = (X_1 + \cdots + X_n)/n$ denotes the sample mean, then by the usual central limit theorem, $\sqrt{n}(\overline{X} - E(X))/\sqrt{\mathrm{Var}(X)}$ approaches the standard normal distribution as $n \to \infty$.

Sometimes one would be interested in the asymptotes of the extreme order statistics $M_n = \max(X_1, \cdots, X_n)$ and $m_n = \min(X_1, \cdots, X_n)$. Here, we determine the max and min domains of attraction of the cumulative distribution function given by (4).

Let $g(t) = (\delta/\alpha)(t-\mu)^{1-\alpha}$. Then,

$$\begin{aligned}
\lim_{t\to\infty} \frac{1 - F\left(t + xg(t)\right)}{1 - F(t)} &= \lim_{t\to\infty} \exp\left\{\frac{1}{\delta}\left[(t-\mu)^\alpha - (t + xg(t) - \mu)^\alpha\right]\right\} \\
&= \lim_{t\to\infty} \exp\left\{\frac{1}{\delta}(t-\mu)^\alpha\left[1 - \left(1 + \frac{xg(t)}{t-\mu}\right)^\alpha\right]\right\} \\
&= \lim_{t\to\infty} \exp\left\{-\frac{\alpha}{\delta}(t-\mu)^\alpha g(t) x\right\} \\
&= \exp(-x),
\end{aligned}$$

for every $x \in (-\infty, \infty)$. So, it follows by Leadbetter *et al.* (1987, Chapter 1) that F belongs to the max domain of attraction of the Gumbel extreme value distribution with

$$\lim_{n\to\infty} \Pr\left\{a_n\left(M_n - b_n\right) \le x\right\} = \exp\left\{-\exp(-x)\right\},$$

for some suitable norming constants $a_n > 0$ and b_n.

The form of the norming constants can also be determined. For instance, using Corollary 1.6.3 in Leadbetter *et al.* (1987), one can see that $b_n = F^{-1}(1 - 1/n)$ and $a_n = (\alpha/\delta)(b_n - \mu)^{\alpha-1}$, where $F^{-1}(\cdot)$ is given by (6).

For the min domain of attraction, we note that

$$\lim_{t\to 0} \frac{F\left(tx + \mu\right)}{F(t + \mu)} = \lim_{t\to 0}\left(\frac{tx}{t}\right)^\alpha = x^\alpha.$$

So, it follows by Leadbetter *et al.* (1987, Chapter 1) that F belongs to the min domain of attraction of the Weibull extreme value distribution.

7. Maximum Likelihood Estimation

Suppose $x_1, x_2, \cdots, x_n$ is a random sample of size n from (5). Then the log-likelihood function for the vector of parameters $(\alpha, \beta, \delta, \mu)$ can be written as

$$\begin{aligned} \ln L(\alpha, \beta, \delta, \mu) = & \; n \ln \alpha - n \ln \delta - n \ln \beta - n \ln \left[1 - 2^{-1/\beta}\right] - \frac{1}{\delta} \sum_{i=1}^{n} (x_i - \mu)^{\alpha} \\ & + (\alpha - 1) \sum_{i=1}^{n} \ln (x_i - \mu) \\ & - \left(\frac{1}{\beta} + 1\right) \sum_{i=1}^{n} \ln \left\{1 + \exp\left[-\frac{1}{\delta} (x_i - \mu)^{\alpha}\right]\right\}. \end{aligned} \tag{16}$$

The first-order partial derivatives of (16) with respect to the four parameters are:

$$\begin{aligned} \frac{\partial \ln L}{\partial \alpha} = & \; \frac{n}{\alpha} - \frac{1}{\delta} \sum_{i=1}^{n} (x_i - \mu)^{\alpha} \ln (x_i - \mu) + \sum_{i=1}^{n} \ln (x_i - \mu) \\ & + \frac{1}{\delta} \left(\frac{1}{\beta} + 1\right) \sum_{i=1}^{n} \frac{\exp\left[-(x_i - \mu)^{\alpha} / \delta\right] (x_i - \mu)^{\alpha} \ln (x_i - \mu)}{1 + \exp\left[-(x_i - \mu)^{\alpha} / \delta\right]}, \end{aligned} \tag{17}$$

$$\frac{\partial \ln L}{\partial \beta} = -\frac{n}{\beta} + \frac{n \ln 2}{\beta^2 \left(2^{1/\beta} - 1\right)} + \frac{1}{\beta^2} \sum_{i=1}^{n} \ln \left\{1 + \exp\left[-\frac{1}{\delta} (x_i - \mu)^{\alpha}\right]\right\}, \tag{18}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \delta} = & -\frac{n}{\delta} + \frac{1}{\delta^2} \sum_{i=1}^{n} (x_i - \mu)^{\alpha} \\ & - \frac{1}{\delta^2} \left(\frac{1}{\beta} + 1\right) \sum_{i=1}^{n} \frac{\exp\left[-(x_i - \mu)^{\alpha} / \delta\right] (x_i - \mu)^{\alpha}}{1 + \exp\left[-(x_i - \mu)^{\alpha} / \delta\right]}, \end{aligned} \tag{19}$$

and

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} = & \; \frac{\alpha}{\delta} \sum_{i=1}^{n} (x_i - \mu)^{\alpha - 1} - (\alpha - 1) \sum_{i=1}^{n} (x_i - \mu)^{-1} \\ & - \frac{\alpha}{\delta} \left(\frac{1}{\beta} + 1\right) \sum_{i=1}^{n} \frac{\exp\left[-(x_i - \mu)^{\alpha} / \delta\right] (x_i - \mu)^{\alpha - 1}}{1 + \exp\left[-(x_i - \mu)^{\alpha} / \delta\right]}. \end{aligned} \tag{20}$$

The maximum likelihood estimators of $(\alpha, \beta, \delta, \mu)$, say $(\widehat{\alpha}, \widehat{\beta}, \widehat{\delta}, \widehat{\mu})$, are the simultaneous solutions of the equations $\partial \ln L / \partial \alpha = 0$, $\partial \ln L / \partial \beta = 0$, $\partial \ln L / \partial \delta = 0$ and $\partial \ln L / \partial \mu = 0$. As $n \to \infty$, $(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta, \widehat{\delta} - \delta, \widehat{\mu} - \mu)$ approaches a multivariate normal vector with zero means and variance-covariance matrix, $-(E\boldsymbol{J})^{-1}$, where

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial \alpha^2} & \frac{\partial^2 \ln L}{\partial \alpha \partial \beta} & \frac{\partial^2 \ln L}{\partial \alpha \partial \delta} & \frac{\partial^2 \ln L}{\partial \alpha \partial \mu} \\ \frac{\partial^2 \ln L}{\partial \beta \partial \alpha} & \frac{\partial^2 \ln L}{\partial \beta^2} & \frac{\partial^2 \ln L}{\partial \beta \partial \delta} & \frac{\partial^2 \ln L}{\partial \beta \partial \mu} \\ \frac{\partial^2 \ln L}{\partial \delta \partial \alpha} & \frac{\partial^2 \ln L}{\partial \delta \partial \beta} & \frac{\partial^2 \ln L}{\partial \delta^2} & \frac{\partial^2 \ln L}{\partial \delta \partial \mu} \\ \frac{\partial^2 \ln L}{\partial \mu \partial \alpha} & \frac{\partial^2 \ln L}{\partial \mu \partial \beta} & \frac{\partial^2 \ln L}{\partial \mu \partial \delta} & \frac{\partial^2 \ln L}{\partial \mu^2} \end{pmatrix}.$$

The matrix, $-E\boldsymbol{J}$, is known as the expected information matrix. The matrix, $-\boldsymbol{J}$, is known as the observed information matrix.

In simulations and real data applications described later on, we maximized the log-likelihood function using the nlm function in the R statistical package (R Development Core Team, 2012). For each maximization, the nlm function was executed for a wide range of initial values. This sometimes resulted in more than one maximum, but at least one maximum was identified each time. In cases of more than one maximum, we took the maximum likelihood estimates to correspond to the largest of the maxima.

In practice, n is finite. The literature (see, for example, Efron and Hinkley, 1978) suggests that it is best to approximate the distribution of $(\widehat{\alpha}-\alpha, \widehat{\beta}-\beta, \widehat{\delta}-\delta, \widehat{\mu}-\mu)$ by a multivariate normal distribution with zero means and variance-covariance matrix given by $-\boldsymbol{J}^{-1}$, inverse of the observed information matrix, with $(\alpha, \beta, \delta, \mu)$ replaced $(\widehat{\alpha}, \widehat{\beta}, \widehat{\delta}, \widehat{\mu})$. So, it is useful to have explicit expressions for the elements of $\boldsymbol{J}$. They are given in Appendix A.

The multivariate normal approximation can be used to construct approximate confidence intervals and confidence regions for the individual parameters and for the hazard and survival functions. A natural question is: how large should n be for the approximation to be good enough? Section 8 gives an answer to this question by means of a simulation study.

8. Simulation Study

Here, we assess the performance of the maximum likelihood estimators given by (18)-(20) with respect to sample size n. The assessment is based on a simulation study:

1. generate ten thousand samples of size n from (5). The inversion method was used to generate samples, i.e., variates of the Chen distribution were generated using (6);

2. compute the maximum likelihood estimates for the ten thousand samples, say $(\widehat{\alpha}_i, \widehat{\beta}_i, \widehat{\delta}_i, \widehat{\mu}_i)$ for $i = 1, 2, \cdots, 10000$;

3. compute the biases and mean squared errors given by

$$\text{bias}_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} \left(\widehat{h}_i - h\right),$$

and

$$\text{MSE}_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} \left(\widehat{h}_i - h\right)^2,$$

for $h = \alpha, \beta, \delta, \mu$.

We repeat these steps for $n = 10, 20, \cdots, 1000$ with $\alpha = 2$, $\beta = 2$, $\delta = 1$ and $\mu = 0$, so computing $\text{bias}_\alpha(n)$, $\text{bias}_\beta(n)$, $\text{bias}_\delta(n)$, $\text{bias}_\mu(n)$ and $\text{MSE}_\alpha(n)$, $\text{MSE}_\beta(n)$, $\text{MSE}_\delta(n)$, $\text{MSE}_\mu(n)$ for $n = 10, 20, \cdots, 1000$.

Figures 4 and 5 show how the four biases and the four mean squared errors vary with respect to n. The broken line in Figure 4 corresponds to the biases being zero. The broken line in Figure 5 corresponds to the mean squared errors being zero. The following observations can be made:

1. the biases for α, β and δ are generally negative;
2. the biases for μ are generally positive;
3. the biases for each parameter generally approach zero as $n \to \infty$;
4. the biases appear smallest for δ;
5. the mean squared errors for each parameter generally decrease to zero as $n \to \infty$;
6. the mean squared errors appear largest for α;
7. the mean squared errors appear smallest for μ for n large enough.

We have presented results for only one choice for $(\alpha, \delta, \delta, \mu)$, namely that $(\alpha, \delta, \delta, \mu) = (2, 2, 1, 0)$. But the results were similar for other choices.

For the real data application presented in Section 9, we have $n = 52$. We see from Figure 4 that the biases of all four of the parameters for $n = 52$ are less than 0.04. We see from Figure 5 that the mean squared errors of all four of the parameters for $n = 52$ are less than 0.1. So, it is reasonable to assume that the normal approximation holds for the data application.

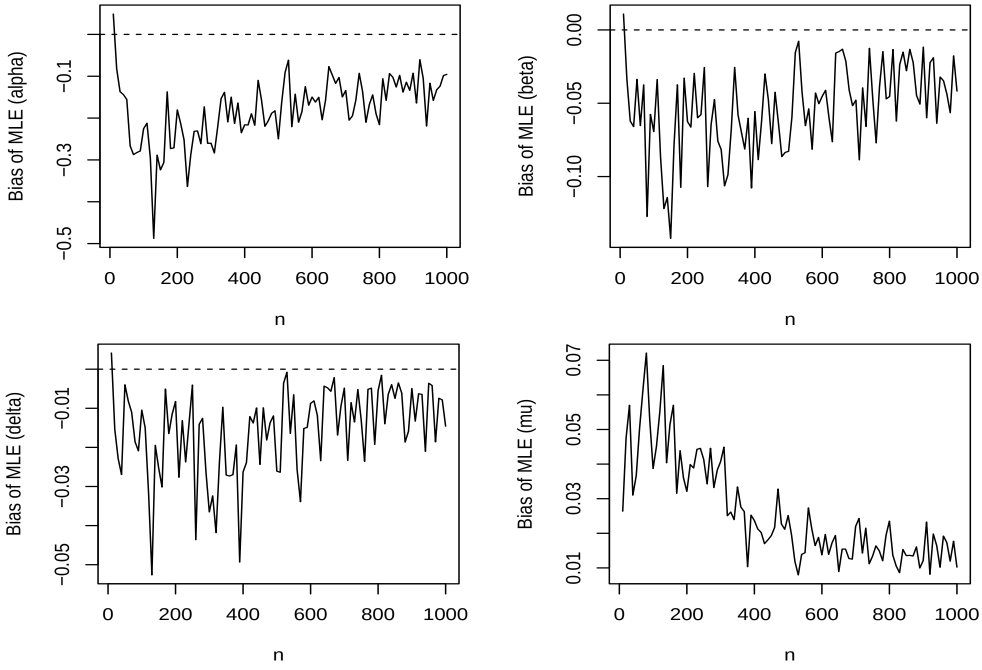

Figure 4: $\text{bias}_\alpha(n)$ (top left), $\text{bias}_\beta(n)$ (top right), $\text{bias}_\delta(n)$ (middle right) and $\text{bias}_\mu(n)$ (bottom left) versus $n = 10, 20, \cdots, 1000$

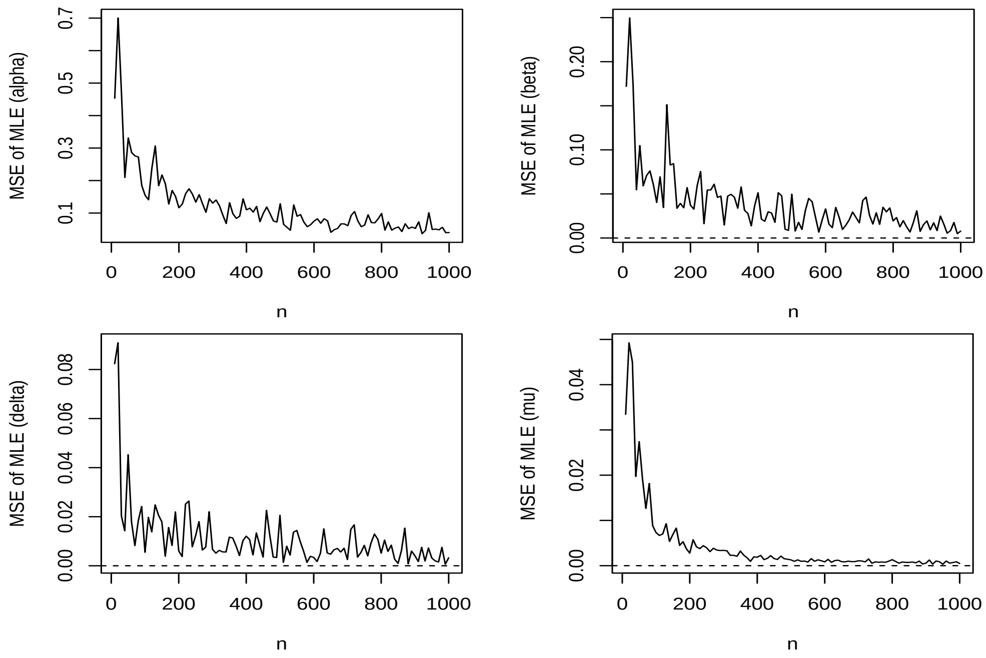

Figure 5: $\text{MSE}_\alpha(n)$ (top left), $\text{MSE}_\beta(n)$ (top right), $\text{MSE}_\delta(n)$ (middle right) and $\text{MSE}_\mu(n)$ (bottom left) versus $n = 10, 20, \cdots, 1000$

9. An Application

Here, we illustrate flexibility of the Chen distribution using the real data set used in Chen *et al.* (2010). The data used by Chen *et al.* (2010) are fifty two ordered annual maximum antecedent rainfall measurements in mm from Maple Ridge in British Columbia, Canada: 264.9, 314.1, 364.6, 379.8, 419.3, 457.4, 459.4, 460.0, 490.3, 490.6, 502.2, 525.2, 526.8, 528.6, 528.6, 537.7, 539.6, 540.8, 551.0, 573.5, 579.2, 588.2, 588.7, 589.7, 592.1, 592.8, 600.8, 604.4, 608.4, 609.8, 619.2, 626.4, 629.4, 636.4, 645.2, 657.6, 663.5, 664.9, 671.7, 673.0, 682.6, 689.8, 698.0, 698.6, 698.8, 703.2, 755.9, 786.0, 787.2, 798.6, 850.4, 895.1.

We fitted the distributions given (1) and (5) to the data. The maximum likelihood procedure described in Section 7 was used for fitting (5). The fitted estimates for (5) were: $\widehat{\alpha} = 1.700$ (0.069), $\widehat{\beta} = 0.055$ (0.045), $\widehat{\delta} = 17125.09$ (10190.82), $\widehat{\mu} = -33.524$ (193.361) with $-\ln L = 324.443$ and AIC $= 656.885$. The fitted estimates for (1) were: $\widehat{\mu} = 552.016$ (19.520), $\widehat{\sigma} = 129.755$ (13.307), $\widehat{\xi} = -0.308$ (0.072) with $-\ln L = 327.932$ and AIC $= 661.864$. The numbers within brackets are standard errors obtained by inverting the observed information matrix, see Section 7.

We can see that the negative log-likelihood values and the AIC values are smaller for the Chen distribution. So, for the data set used in Chen *et al.* (2010), the Chen distribution provides a better fit. This is confirmed by the probability-probability plots, quantile-quantile plots and density plots shown in Figures 6 to 8. The points in Figures 6 and 7 are closer to the diagonal lines for the Chen distribution. The fitted probability density function for the Chen distribution appears to better capture the histogram in Figure 8.

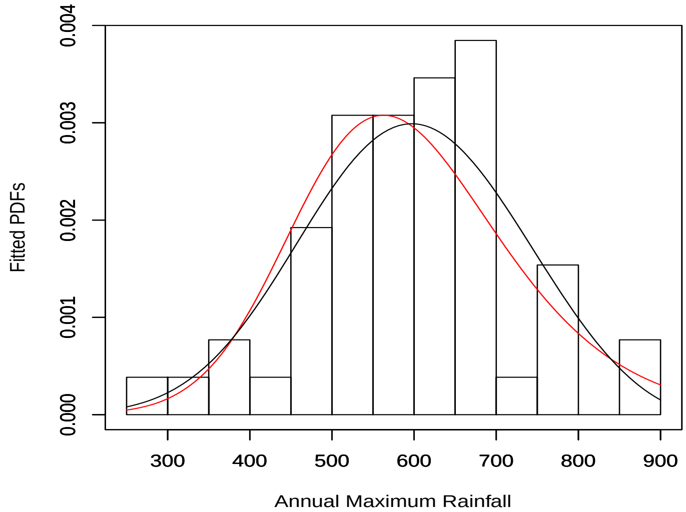

Figure 6: Probability plots for the fits of (1) (in red) and (5) (in black) for annual maximum rainfall from Maple Ridge in British Columbia

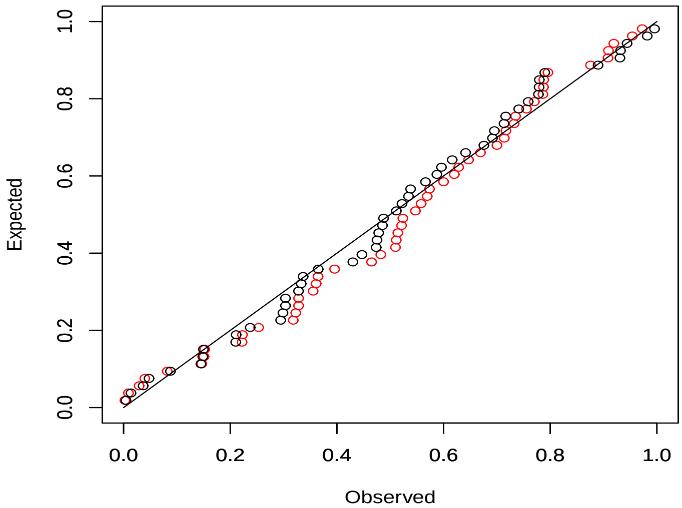

Figure 7: Quantile plots for the fits of (1) (in red) and (5) (in black) for annual maximum rainfall from Maple Ridge in British Columbia

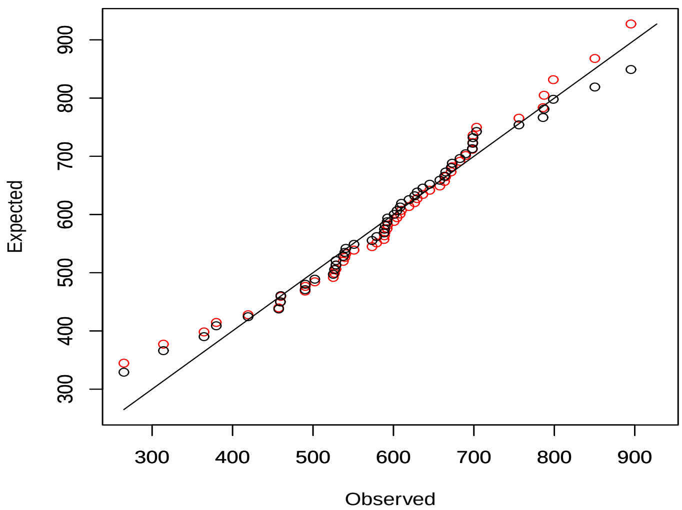

Figure 8: Density plots for the fits of (1) (in red) and (5) (in black) for annual maximum rainfall from Maple Ridge in British Columbia

Furthermore, chi-square goodness of fit tests gave the p-values of 0.039 and 0.071 for (1) and (5), respectively, suggesting that (5) provides the only adequate fit.

Appendix A

Here, we give explicit expressions for the elements of $\boldsymbol{J}$ defined in Section 7:

$$J_{11} = -\frac{n}{\alpha^2} - \frac{1}{\delta}\sum_{i=1}^{n}(x_i-\mu)^{\alpha}\ln^2(x_i-\mu) + \frac{1}{\delta}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha}\ln^2(x_i-\mu)}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}$$
$$-\frac{1}{\delta^2}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{\exp\left[(x_i-\mu)^{\alpha}/\delta\right](x_i-\mu)^{2\alpha}\ln^2(x_i-\mu)}{\left\{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]\right\}^2},$$

$$J_{12} = -\frac{1}{\delta\beta^2}\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha}\ln(x_i-\mu)}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]},$$

$$J_{13} = \frac{1}{\delta^2}\sum_{i=1}^{n}(x_i-\mu)^{\alpha}\ln(x_i-\mu) - \frac{1}{\delta^2}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha}\ln(x_i-\mu)}{1+\exp\left[(x_i-\mu)^{\alpha}/\alpha\right]}$$
$$+\frac{1}{\delta^3}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{2\alpha}\ln(x_i-\mu)\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}{\left\{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]\right\}^2},$$

$$J_{14} = \frac{\alpha}{\delta}\sum_{i=1}^{n}(x_i-\mu)^{\alpha-1}\ln(x_i-\mu) + \frac{1}{\delta}\sum_{i=1}^{n}(x_i-\mu)^{\alpha-1} - \sum_{i=1}^{n}(x_i-\mu)^{-1}$$
$$-\frac{1}{\delta}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha-1}\left[1+\alpha\ln(x_i-\mu)\right]}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}$$
$$+\frac{\alpha}{\delta^2}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{2\alpha-1}\ln(x_i-\mu)\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}{\left\{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]\right\}^2},$$

$$J_{22} = \frac{n}{\beta^2} - \frac{2n\ln 2}{\beta^3\left(2^{1/\beta}-1\right)} + \frac{n(\ln 2)^2 2^{1/\beta}}{\beta^4\left(2^{1/\beta}-1\right)^2} - \frac{2}{\beta^3}\sum_{i=1}^{n}\ln\left\{1+\exp\left[-\frac{1}{\delta}(x_i-\mu)^{\alpha}\right]\right\},$$

$$J_{23} = \frac{1}{\beta^2\delta^2}\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha}}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]},$$

$$J_{24} = \frac{\alpha}{\beta^2\delta}\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha-1}}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]},$$

$$J_{33} = \frac{n}{\delta^2} - \frac{2}{\delta^3}\sum_{i=1}^{n}(x_i-\mu)^{\alpha} + \frac{2}{\delta^3}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha}}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}$$
$$-\frac{1}{\delta^4}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{2\alpha}\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}{\left\{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]\right\}^2},$$

$$J_{34} = -\frac{\alpha}{\delta^2}\sum_{i=1}^{n}(x_i-\mu)^{\alpha-1} + \frac{\alpha}{\delta^2}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha-1}}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]}$$
$$-\frac{\alpha}{\delta^3}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{\exp\left[(x_i-\mu)^{\alpha}/\delta\right](x_i-\mu)^{2\alpha-1}}{\left\{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]\right\}^2},$$

and

$$J_{44} = -\frac{\alpha(\alpha-1)}{\delta}\sum_{i=1}^{n}(x_i-\mu)^{\alpha-2} - (\alpha-1)\sum_{i=1}^{n}(x_i-\mu)^{-2} + \frac{\alpha(\alpha-1)}{\delta}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{(x_i-\mu)^{\alpha-2}}{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]} - \frac{\alpha^2}{\delta^2}\left(\frac{1}{\beta}+1\right)\sum_{i=1}^{n}\frac{\exp\left[1\,(x_i-\mu)^{\alpha}/\delta\right](x_i-\mu)^{2\alpha-2}}{\left\{1+\exp\left[(x_i-\mu)^{\alpha}/\delta\right]\right\}^2}.$$

Explicit expressions for the remaining elements of $\boldsymbol{J}$ follow by symmetry.

Acknowledgements

The authors would like to thank the Editor and the referee for careful reading and for their comments which greatly improved the paper.

References

Castillo, E., Hadi, A. S., Balakrishnan, N. and Sarabia, J. M. (2005). *Extreme Value and Related Models with Applications in Engineering and Science*. Wiley, New Jersey.

Chen, G., Bunce, C. and Jiang, W. (2010). A new distribution for extreme value analysis. In *Proceedings of the International Conference on Computational Intelligence and Software Engineering*, 1-4. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5677106.

Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimation: observed versus expected Fisher information. *Biometrika* **65**, 457-482.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Springer, Berlin.

Hosking, J. R. M. (1990). *L*-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B* **52**, 105-124.

Hosking, J. R. M. (1994). The four-parameter kappa distribution. *IBM Journal of Research and Development* **38**, 251-258.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1987). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

Mielke, P. W. (1973). Another family of distributions for describing and analyzing precipitation data. *Journal of Applied Meteorology* **12**, 275-280.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/.

Resnick, S. I. (2008). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

Silva, R. B., Barreto-Souza, W. and Cordeiro, G. M. (2010). A new distribution with decreasing, increasing and upside-down bathtub failure rate. *Computational Statistics and Data Analysis* **54**, 935-944.

Singh, H. and Misra, N. (1994). On redundancy allocations in systems. *Journal of Applied Probability* **31**, 1004-1014.

Received April 18, 2013; accepted July 1, 2013.

Saralees Nadarajah
School of Mathematics
University of Manchester
Manchester M13 9PL, UK
Saralees.Nadarajah@manchester.ac.uk

Sumaya Eljabri
School of Mathematics
University of Manchester
Manchester M13 9PL, UK