2.1 Introduction

In probability and statistics, the Gumbel distribution is used to model the distribution of the maximum (or the minimum) of a number of samples of various distributions. The importance of the Gumbel distribution in practice is due to its extreme value behavior. It has been applied either as the parent distribution or as an asymptotic approximation, to describe extreme wind speeds, sea wave heights, floods, rainfall, age at death, minimum temperature, rainfall during droughts, electrical strength of materials, air pollution problems, geological problems, naval engineering etc. A generalization of the Gumbel distribution is discussed by Adeyemi and Ojo (2003). Inference for $P(Y < X)$ in exponentiated Gumbel distribution is studied by Kakade et al. (2008) and exponentiated Gumbel distribution for estimation of return levels of significant wave height is considered by Persson and Ryden (2010). Estimation of the extreme value type 1 distribution by the method of L-Q moments is discussed
by Shabri and Jemain (2009).

A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families is introduced by Marshall and Olkin (1997). For a random variable with a distribution function \( F(x) \) and survival function \( F^{-}(x) \), we can obtain a new family of distribution functions called univariate Marshall-Olkin family having cumulative distribution function \( G(x) \) given by

\[
G(x) = \frac{F(x)}{\alpha + (1 - \alpha)F(x)}; -\infty < x < \infty; 0 < \alpha < \infty.
\]

Then the corresponding survival function is

\[
\bar{G}(x) = \frac{\alpha F(x)}{1 - (1 - \alpha)F(x)}; -\infty < x < \infty; 0 < \alpha < \infty.
\]

This new family involves an additional parameter \( \alpha \). Jose et al. (2010) discussed Marshall Olkin q-Weibull distribution and Max-Min processes. Alice and Jose (2005 a, b) introduced Marshall Olkin distributions with semi Weibull and Logistic marginals. Bivariate Marshall-Olkin exponential minification processes are discussed by Ristic et al. (2008).

A minification processes of the first order is given by

\[
X_n = K \min(X_{n-1}, \varepsilon_n), n \geq 1,
\]

where \( \{\varepsilon_n, n \geq 1\} \) where \( k > 1 \) is a constant, and \( \{\varepsilon_n\} \) is an innovation process of independently and identically distributed random variables chosen to ensure that \( \{X_n\} \) is a stationary markov process with marginal distribution function \( F_X(x) \). Because of the structure of (2.1.1) the process \( \{X_n\} \) is called a minification processes (See Lewis and McKenzie (1991)). Alice and Jose (2004) developed a bivariate minification processes with semi Pareto marginal distribution. Jose et al. (2011) introduced a minification processes with Marshall Olkin bivariate Weibull distribution. Minification processes with discrete marginals is discussed by Kalamkar (1995).
In this chapter Gumbel distribution is discussed in section 2.2. In section 2.3, Marshall-Olkin Gumbel maximum distribution and its properties are discussed. In section 2.4, we consider Marshall-Olkin Gumbel minimum distribution. Some properties are discussed in section 2.5. Minification Processes with Marshall-Olkin Gumbel marginal distribution are also developed and studied in section 2.6. Estimation of reliability when X and Y are independent Marshall Olkin Gumbel maximum distribution is done in section 2.7. Data analysis and modelling with respect to a real data on daily discharge of river water is carried out in section 2.8. Conclusions are given in section 2.9.

2.2 Gumbel distribution

Consider a random sample $X_1, X_2, \ldots, X_n$ taken from a population with common distribution function $F$. The expression for the distribution of $X_{\text{max}}$ in terms of the population distribution function $F$ is given in (2.2.1) as

$$P(X \leq x) = P(X_1 \leq x, \ldots, X_n \leq x) = P(X_1 \leq x) \cdot \ldots \cdot P(X_n \leq x) = \{F(x)\}^n$$

The non-normality of the distribution of $X_{\text{max}}$ is not surprising. In fact, careful mathematical treatment of equation (2.2.1) suggests that for large $n$ the distribution function of $X_{\text{max}}$ should be approximately what is known as a Gumbel distribution. This has distribution function

$$F(x) = e^{-e^{-\frac{x-\lambda}{\delta}}}$$

The corresponding pdf is

$$f(x) = \frac{1}{\delta} e^{-e^{-\frac{x-\lambda}{\delta}}} e^{-\frac{x-\lambda}{\delta}} \quad x \in \mathbb{R}, \quad \delta > 0, \quad \lambda \in \mathbb{R},$$

where $\lambda$ and $\delta$ are constants known as the location and scale parameters. We consider now the maximum likelihood estimation of the unknown parameters $\lambda$ and $\delta$. We estimate
Table 2.1: The population parameters of the Gumbel distribution are summarized.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>maxima</th>
<th>minima</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>$\lambda + 0.57772\delta$</td>
<td>$\lambda - 0.57772\delta$</td>
</tr>
<tr>
<td>Median</td>
<td>$\lambda + 0.3665\delta$</td>
<td>$\lambda - 0.3665\delta$</td>
</tr>
<tr>
<td>Mode</td>
<td>$\lambda$</td>
<td>$\lambda$</td>
</tr>
<tr>
<td>Variance</td>
<td>$\frac{\pi^2\delta^2}{6}$</td>
<td>$\frac{\pi^2\delta^2}{6}$</td>
</tr>
<tr>
<td>Kurtosis coeff.</td>
<td>1.1396</td>
<td>-1.1396</td>
</tr>
</tbody>
</table>

the unknown parameters by using the maximum likelihood function. The log-likelihood function is

$$\log L = -n \log \delta - e^{-\sum_{i=1}^{n} \frac{(x_i - \lambda)}{\delta}} - \sum_{i=1}^{n} \frac{(x_i - \lambda)}{\delta}$$

The partial derivatives are

$$\frac{\partial \log L}{\partial \lambda} = e^{-\sum_{i=1}^{n} \frac{(x_i - \lambda)}{\delta}} - n$$

and

$$\frac{\partial \log L}{\partial \delta} = \sum_{i=1}^{n} (x_i - \lambda)(1 - e^{-\frac{x_i - \lambda}{\delta}}) - n\delta$$

Solving the equations $\frac{\partial \log L}{\partial \lambda} = 0$ and $\frac{\partial \log L}{\partial \delta} = 0$ we obtain the maximum likelihood estimates of the unknown parameters.

2.3 Marshall-Olkin Gumbel Maximum Family

Consider a survival function $F$. Then a new family of survival functions introduced by Marshall-Olkin (1997) is constructed as

$$\overline{G}(x; \alpha) = \frac{\alpha F(x)}{1 - (1 - \alpha)F(x)}, \ x \in \mathbb{R}, \ \alpha > 0.$$  \hspace{1cm} (2.3.1)

Clearly when $\alpha = 1$ we get $\overline{G} = F$. Whenever $F$ has a density, the survival function $\overline{G}$ given by (2.3.1) have easily computed densities. In particular, if $F$ has a density $f$ and its
hazard rate \( r_F \), then \( G \) has the density \( g \) given by

\[
g(x; \alpha) = \frac{\alpha f(x)}{(1 - (1 - \alpha)F(x))^2}, \quad x \in R, \quad \alpha > 0
\]

and hazard rate

\[
r(x; \alpha) = \frac{r_F(x)}{1 - (1 - \alpha)F(x)}, \quad x \in R, \quad \alpha > 0.
\]

Consider the Gumbel maximum distribution with survival function (see Castillo (1988))

\[
F(x) = 1 - e^{-e^{(-x - \lambda)/\delta}}, \quad x \in R, \quad \delta > 0, \quad \lambda \in R,
\]

where \( \lambda \) and \( \delta \) are constants known as the location and scale parameters. Substituting this in (2.3.1) we get a new family of distributions, which we shall refer to as Marshall-Olkin Gumbel Maximum (MO-GUMX) family, whose survival function is given by

\[
G(x) = \frac{\alpha}{\alpha + (1 - \alpha)e^{-e^{(-x - \lambda)/\delta}}}, \quad x \in R, \quad \delta, \alpha > 0, \quad \lambda \in R.
\]

The probability density function is

\[
g(x) = \frac{\alpha e^{-\frac{x - \lambda}{\delta}} e^{-e^{-\frac{x - \lambda}{\delta}}}}{\delta \left(\alpha + (1 - \alpha)e^{-e^{-\frac{x - \lambda}{\delta}}}\right)^2}, \quad x \in R, \quad \delta, \alpha > 0, \quad \lambda \in R.
\]

The probability density function \( g(x) \) has a unique mode at \( x = x_0 \), where \( x_0 \) is the solution of the equation

\[
-1 + \alpha - (1 - \alpha)s(x) - \alpha e^{s(x)} + \alpha s(x)e^{s(x)} = 0,
\]

and \( s(x) = e^{-\frac{x - \lambda}{\delta}} \).
Furthermore, \( g(-\infty) = g(\infty) = 0 \). The hazard rate function is

\[
r(x) = \frac{e^{-\frac{x}{\delta}} e^{-\frac{e^{-\frac{x}{\lambda}}}{\delta}}}{\delta \left( \alpha + (1 - \alpha)e^{-\frac{e^{-\frac{x}{\lambda}}}{\delta}} \right) \left( 1 - e^{-\frac{e^{-\frac{x}{\lambda}}}{\delta}} \right)}, \quad x \in R, \quad \delta, \alpha > 0, \ \lambda \in R.
\]

If \( 0 < \alpha < 0.5 \), then the hazard rate function \( r(x) \) has a maximum at \( x = x_1 \), where \( x_1 \) is the solution of the equation

\[
1 - \alpha + (1 - \alpha)s(x) + (2\alpha - 1)e^{s(x)} - \alpha e^{2s(x)} + \alpha s(x) e^{2s(x)} = 0.
\]

Furthermore, \( r(-\infty) = 0 \) and \( r(\infty) = 1/\delta \). If \( \alpha > 0.5 \), then the hazard rate function \( r(x) \) is an increasing function with \( r(-\infty) = 0 \) and \( r(\infty) = 1/\delta \).

Let us consider the \( n^{th} \) moment of the MO-GUMX distribution. We will consider MO-GUMX(0, 1, \( \alpha \)) distribution, since if \( X \overset{d}{=} \text{MO-GUMX}(\lambda, \delta, \alpha) \), then \( Y = (X - \lambda)/\delta \overset{d}{=} \text{MO-GUMX}(0, 1, \alpha) \). Let \( Y \overset{d}{=} \text{MO-GUMX}(0, 1, \alpha) \). Then the \( n^{th} \) moment of the random variable \( Y \) can be written as

\[
E(Y^n) = \alpha \int_{-\infty}^{\infty} \frac{y^n e^{-y} e^{-\frac{e^{-y}}{\delta}} dy}{(\alpha + (1 - \alpha)e^{-\frac{e^{-y}}{\delta}})^2} = \alpha (-1)^n \int_{0}^{\infty} \frac{(\log u)^n e^{-u} du}{(\alpha + (1 - \alpha)e^{-u})^2}.
\]

Using the expansions

\[
\frac{1}{(\alpha + (1 - \alpha)e^{-u})^2} = \begin{cases} 
\sum_{i=1}^{\infty} i(1 - \alpha)^{i-1} \sum_{j=0}^{i-1} (-1)^j (i-1)^j e^{-ju}, & 0 < \alpha < 1 \\
\frac{1}{\alpha^2} \sum_{i=1}^{\infty} i \left( \frac{\alpha - 1}{\alpha} \right)^{i-1} e^{-(i-1)u}, & \alpha \geq 1
\end{cases}
\]

and from Prudnikov et al. (1986)

\[
\int_{0}^{\infty} (\log u)^n e^{-ju} du = \left. \frac{\partial}{\partial a} \right|_{a=1} \left( \Gamma(a) \right)^n \left( \frac{\Gamma(a)}{j^a} \right)
\]
we obtain that the $n^{th}$ moment of $Y \overset{d}{=} \text{MO-GUMX} (0, 1, \alpha)$ is

$$E(Y^n) = \begin{cases} \alpha^{-1} \sum_{i=1}^{\infty} i(1 - \alpha)^{i-1} \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} \left( \frac{\Gamma(a)}{(j+1)^a} \right) \Bigg|_{a=1}, & 0 < \alpha < 1 \\ \frac{(-1)^n}{\alpha} \sum_{i=1}^{\infty} i \left( \frac{a-1}{\alpha} \right)^{i-1} \left( \frac{\partial}{\partial a} \right)^n \left( \frac{\Gamma(a)}{\alpha^a} \right) \Bigg|_{a=1}, & \alpha \geq 1. \end{cases}$$

We consider now the maximum likelihood estimation of the unknown parameters $\lambda$, $\delta$ and $\alpha$. We estimate the unknown parameters by using the maximum likelihood function. The log-likelihood function is

$$\log L = n \log \alpha - \frac{1}{\delta} \sum_{i=1}^{n} (x_i - \lambda) - \sum_{i=1}^{n} s(x_i) - n \log \delta - 2 \sum_{i=1}^{n} \log \left( \alpha + (1 - \alpha)e^{-s(x_i)} \right).$$

The first partial derivatives are

$$\frac{\partial \log L}{\partial \lambda} = n - \frac{1}{\delta} \sum_{i=1}^{n} s(x_i) + \frac{2(1 - \alpha)}{\delta} \sum_{i=1}^{n} \frac{s(x_i)e^{-s(x_i)}}{\alpha + (1 - \alpha)e^{-s(x_i)}}$$

$$\frac{\partial \log L}{\partial \delta} = \frac{1}{\delta^2} \sum_{i=1}^{n} (x_i - \lambda) - \frac{1}{\delta^2} \sum_{i=1}^{n} (x_i - \lambda)s(x_i) - \frac{n}{\delta}$$

$$+ \frac{2(1 - \alpha)}{\delta^2} \sum_{i=1}^{n} \frac{(x_i - \lambda)s(x_i)e^{-s(x_i)}}{\alpha + (1 - \alpha)e^{-s(x_i)}}$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} - 2 \sum_{i=1}^{n} \frac{1 - e^{-s(x_i)}}{\alpha + (1 - \alpha)e^{-s(x_i)}}.$$
2.4 Marshall-Olkin Gumbel Minimum Distributions

In this section we consider Gumbel minimum distribution with survival function (See Castillo (1988))

\[ F(x) = e^{-e^{-\frac{\lambda - x}{\delta}}}, \quad x \in \mathbb{R}, \delta > 0, \lambda \in \mathbb{R}, \]

where \( \lambda \) and \( \delta \) are constants known as the location and scale parameters. Substituting this in (2.3.1), we get a new family of distributions, which we shall refer to as Marshall-Olkin Gumbel Minimum (MO-GUMN) family, whose survival function is given by

\[ G(x) = \frac{\alpha e^{-e^{-\frac{\lambda - x}{\delta}}}}{1 - (1 - \alpha) e^{-e^{-\frac{\lambda - x}{\delta}}}}, \quad x \in \mathbb{R}, \delta, \alpha > 0, \lambda \in \mathbb{R}. \]

The probability density function is

\[ g(x) = \frac{\alpha e^{-\frac{\lambda - x}{\delta}} e^{-e^{-\frac{\lambda - x}{\delta}}}}{\delta \left(1 - (1 - \alpha) e^{-e^{-\frac{\lambda - x}{\delta}}}\right)^2}, \quad x \in \mathbb{R}, \delta, \alpha > 0, \lambda \in \mathbb{R}. \]

The probability density function \( g(x) \) has a unique mode at \( x = x_0 \), where \( x_0 \) is the solution of the equation

\[ 1 - s_1(x) - (1 - \alpha)e^{-s_1(x)} - (1 - \alpha)s_1(x)e^{-s_1(x)} = 0 \]

and \( s_1(x) = e^{-\frac{\lambda - x}{\delta}} \). Furthermore, we have that \( g(-\infty) = g(\infty) = 0 \). The hazard rate function is given by

\[ r(x) = \frac{e^{-\frac{\lambda - x}{\delta}}}{\delta \left(1 - (1 - \alpha) e^{-e^{-\frac{\lambda - x}{\delta}}} \right)}, \quad x \in \mathbb{R}, \delta, \alpha > 0, \lambda \in \mathbb{R}. \]

The hazard rate function \( r(x) \) is an increasing function with \( r(-\infty) = 0 \) and \( r(\infty) = \infty \). Let \( Y \overset{d}{=} \text{MO-GUMN} \left(0, 1, \alpha\right) \). Then the \( n^{th} \) moment of the random variable \( Y \) can be
written as

\[ E(Y^n) = \alpha \int_{-\infty}^{\infty} \frac{y^n e^y e^{-\alpha y}}{(1 - (1 - \alpha)e^{-\alpha y})^2} dy = \alpha \int_{0}^{\infty} \frac{(\log u)^n e^{-\alpha u} du}{(1 - (1 - \alpha)e^{-\alpha u})^2}. \]

Using the expansions

\[ \frac{1}{(1 - (1 - \alpha)e^{-\alpha u})^2} = \begin{cases} 
\sum_{i=1}^{\infty} i(1 - \alpha)^{i-1} e^{-(i-1)u}, & 0 < \alpha < 1 \\
\frac{1}{\alpha^2} \sum_{i=1}^{\infty} i \left( \frac{\alpha-1}{\alpha} \right)^{i-1} \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} e^{-ju}, & \alpha \geq 1 
\end{cases} \quad (2.4.1) \]

and from Prudnikov et al. (1986), we obtain that the \( n^{th} \) moment of \( Y \sim \text{MO-GUMN} (0, 1, \alpha) \) as

\[ E(Y^n) = \begin{cases} 
\alpha \sum_{i=1}^{\infty} i(1 - \alpha)^{i-1} \left( \frac{\Gamma(a)}{a^n} \right) \bigg|_{a=1}, & 0 < \alpha < 1 \\
\frac{1}{\alpha} \sum_{i=1}^{\infty} i \left( \frac{\alpha-1}{\alpha} \right)^{i-1} \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} \left( \frac{\Gamma(a)}{(j+1)\alpha} \right) \bigg|_{a=1}, & \alpha \geq 1 
\end{cases} \]

Consider now the maximum likelihood estimation of the unknown parameters. The log-likelihood function is

\[ \log L = n \log \alpha - \frac{1}{\delta} \sum_{i=1}^{n} (\lambda - x_i) - \sum_{i=1}^{n} s_1(x_i) - n \log \delta - 2 \sum_{i=1}^{n} \log(1 - (1 - \alpha)e^{-s_1(x_i)}). \]

The normal equations are

\[ \frac{\partial \log L}{\partial \lambda} = -\frac{n}{\delta} + \frac{1}{\delta} \sum_{i=1}^{n} s_1(x_i) + \frac{2(1 - \alpha)}{\delta} \sum_{i=1}^{n} s_1(x_i)e^{-s_1(x_i)} = 0 \]

\[ \frac{\partial \log L}{\partial \delta} = \frac{1}{\delta^2} \sum_{i=1}^{n} (\lambda - x_i) - \frac{1}{\delta^2} \sum_{i=1}^{n} (\lambda - x_i)s_1(x_i) - \frac{n}{\delta} - \frac{2(1 - \alpha)}{\delta^2} \sum_{i=1}^{n} \frac{(\lambda - x_i)s_1(x_i)e^{-s_1(x_i)}}{1 - (1 - \alpha)e^{-s_1(x_i)}} = 0 \]

\[ \frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} - 2 \sum_{i=1}^{n} \frac{e^{-s_1(x_i)}}{1 - (1 - \alpha)e^{-s_1(x_i)}} = 0. \]
Graphical representation of probability density function and hazard rate function of Marshall-Olkin Gumbel maximum and Marshall-Olkin Gumbel minimum distributions are given in Fig 2.1a, b, c, d.

2.5 Characteristic Properties

In this section we consider some properties of the distribution.

**Definition 2.5.1.** Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with distribution $F$ in the family (2.3.1) and suppose $N$ is independent of the $X_i$’s with a geometric($p$) distribution such that

$$P(N = n) = p(1 - p)^{n-1}, \ n = 1, 2, \ldots$$

Let $U_N = \min(X_1, X_2, \ldots, X_N)$ and $V_N = \max(X_1, X_2, \ldots, X_N)$. If $F \in \Phi$ implies that the distribution of $U(V)$ is in $\Phi$, then $\Phi$ is said to be geometric minimum stable (geometric maximum stable). If $\Phi$ is both geometric minimum and geometric maximum stable, then $\Phi$ is said to be geometric extreme stable.

**Theorem 2.5.1.** Marshall-Olkin Gumbel maximum distribution is geometric extreme stable.

**Proof.** Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables. Suppose $N$ is independent of the $X_i$’s with a geometric($p$) distribution and let $U_N = \min(X_1, X_2, \ldots, X_N)$ and $V_N = \max(X_1, X_2, \ldots, X_N)$. Then we obtain that

$$\overline{G}(x) = P(U_N > x) = \sum_{n=1}^{\infty} \overline{F}^n(x)(1 - p)^{n-1}p = \frac{p\overline{F}(x)}{1 - (1 - p)\overline{F}(x)}.$$

Suppose that $\overline{F}$ is the survival function of Marshall-Olkin Gumbel maximum family. Then

$$\overline{G}(x) = \frac{\alpha p(1 - e^{-s(x)})}{\alpha p + (1 - \alpha p)e^{-s(x)}}$$
**Figure 2.1**: Graphical representation of probability density function and hazard rate function of Marshall-Olkin Gumbel maximum and Marshall-Olkin Gumbel minimum distributions.

Fig 2.1a. pdf MO-GUMX when $\lambda = 7, \delta = 1, \alpha = .9$

Fig 2.1b. pdf MO-GUMN when $\lambda = 2, \delta = 2, \alpha = 4$

Fig 2.1c. Hazard MO-GUMX when $\lambda = 1, \delta = 1, \alpha = .6$

Fig 2.1d. Hazard MO-GUMN when $\lambda = 1, \delta = 1, \alpha = .6$
and it follows that $U_N$ is geometric minimum stable. Let $V_N = \max(X_1, X_2, \ldots, X_N)$. Then

$$H(x) = P(V_N < x) = \sum_{n=1}^{\infty} F^n(x)(1 - p)^{n-1} p = \frac{pF(x)}{1 - (1 - p)F(x)},$$

so that $H(x) = \frac{F(x)}{p(1 - p)F(x)}$, $x \in R$. If we suppose that $F$ is the survival function of MO-GUMX distribution, then it follows that

$$H(x) = \frac{\theta(1 - e^{-s(x)})}{\theta + (1 - \theta)e^{-s(x)}},$$

where $\theta = \frac{\alpha}{p}$. Hence $V_N$ is geometric maximum stable. Thus the family of MO-GUMX family of distributions is geometric extreme stable.

**Theorem 2.5.2.** Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed random variables with common survival function $F(x)$ and $N$ be a geometric random variable with parameter $p$, which is independent of $\{X_i\}$ for all $i \geq 1$. Let $U_N = \min_{1 \leq i \leq N} X_i$. Then $\{U_N\}$ is distributed as MO-GUMX if and only if $\{X_i\}$ is distributed as GUMX.

**Proof.** Consider $H(x) = P(U_N > x)$, suppose that $F$ is the survival function of the Gumbel maximum family. Then

$$H(x) = \frac{p(1 - e^{-s(x)})}{1 - (1 - p)(1 - e^{-s(x)})},$$

which is the survival function of MO-GUMX. This proves the sufficiency part of the theorem. Conversely suppose

$$H(x) = \frac{p(1 - e^{-s(x)})}{1 - (1 - p)(1 - e^{-s(x)})}.$$

Then we get

$$F(x) = 1 - e^{-e^{-\frac{x}{\lambda}}},$$

which is the survival function of Gumbel maximum.
Theorem 2.5.3. Marshall-Olkin Gumbel minimum distribution is geometric extreme stable.

Proof is similar to Theorem (2.5.1.)

Theorem 2.5.4. Let \( \{X_i, i \geq 1\} \) be a sequence of independent and identically distributed random variables with common survival function \( \overline{F}(x) \) and \( N \) be a geometric random variable with parameter \( p \), which is independent of \( \{X_i\} \) for all \( i \geq 1 \). Let \( U_N = \min_{1 \leq i \leq N} X_i \). Then \( \{U_N\} \) is distributed as MO-GUMN if and only if \( \{X_i\} \) is distributed as GUMN.

Proof is similar to Theorem (2.5.2.)

2.6 Applications in Time Series Modeling

In this section we consider autoregressive minification processes of order 1 and order \( k \).

Theorem 2.6.1. Consider an AR (1) structure given by

\[
X_n = \begin{cases} 
\varepsilon_n & \text{with probability } p \\
\min(X_{n-1}, \varepsilon_n) & \text{with probability } 1 - p
\end{cases}
\]  

(2.6.1)

where \( 0 \leq p \leq 1 \), \( \{\varepsilon_n\} \) is a sequence of independent and identically distributed random variables independent of \( \{X_{n-1}, X_{n-2}, \ldots\} \). Then \( \{X_n\} \) is a stationary Markovian AR(1) process with MO-GUMX marginals if and only if \( \{\varepsilon_n\} \) is distributed as GUMX distribution.

Proof. From (2.6.1) it follows that

\[
\overline{F}_{X_n}(x) = p\overline{F}_{\varepsilon_n}(x) + (1 - p)\overline{F}_{X_{n-1}}(x)\overline{F}_{\varepsilon_n}(x).
\]

Under stationary equilibrium, this gives

\[
\overline{F}_{X}(x) = p\overline{F}_{\varepsilon}(x)/[1 - (1 - p)\overline{F}_{\varepsilon}(x)]
\]
If we take 

\[ F_\varepsilon(x) = 1 - e^{-e^{-\frac{x}{\lambda}}}, \]

then we obtain that \( F_X \) is the survival function of MO-GUMX. Conversely, if we take that 

\[ F_{X_n}(x) = \frac{p \left( 1 - e^{-e^{-\frac{x}{\lambda}}} \right)}{p + (1 - p)e^{-e^{-\frac{x}{\lambda}}}}, \]

then it is easy to show that \( F_{\varepsilon_n}(x) \) is distributed as Gumbel maximum and the process is stationary.

**Theorem 2.6.2.** Consider an autoregressive minification process \( X_n \) of order \( k \) with structure

\[
X_n = \begin{cases} 
\varepsilon_n & \text{with probability } p_0 \\
\min(X_{n-1}, \varepsilon_n) & \text{with probability } p_1 \\
\min(X_{n-2}, \varepsilon_n) & \text{with probability } p_2 \\
\vdots & \vdots \\
\min(X_{n-k}, \varepsilon_n) & \text{with probability } p_k
\end{cases}
\]

(2.6.2)

where \( 0 < p_i < 1 \), \( p_1 + p_2 + \cdots + p_k = 1 - p_0 \). Then \( \{X_n\} \) has stationary marginal distribution as MO-GUMX if and only if \( \{\varepsilon_n\} \) is distributed as GUMX.

As in the case of GUMX distribution we can establish the following theorems.

**Theorem 2.6.3.** Consider an AR(1) structure given by (2.6.1), where \( \{\varepsilon_n\} \) is a sequence of independent and identically distributed random variables independent of \( \{X_{n-1}, X_{n-2}, \ldots\} \). Then \( \{X_n\} \) is stationary Markovian AR(1) process with MO-GUMN marginals if and only if \( \{\varepsilon_n\} \) is distributed as GUMX distribution.

Proof is similar to Theorem (2.6.1).

**Theorem 2.6.4.** Consider an autoregressive minification process \( X_n \) of order \( k \) with structure (2.6.3). Then \( \{X_n\} \) has stationary marginal distribution as MO-GUMN if and only if \( \{\varepsilon_n\} \) is distributed as GUMN.
2.7 Estimation of reliability

Let X and Y be two independent random variables following Marshall Olkin Gumbel maximum distribution with parameters \( \alpha_1, \lambda, \delta \) and \( \alpha_2, \lambda, \delta \) respectively. Then according to Gupta et al. (2009) the reliability of the system given by \( P(X < Y) \) where X is the stress and Y is the strength is given by

\[
R = P(X < Y) = \int_{-\infty}^{\infty} P(Y > X/X = x)g_X(x)dx
\]

\[
= \int_{-\infty}^{\infty} \frac{\alpha_1 \left( 1 - e^{-\frac{x - \lambda}{\delta}} \right)}{\alpha_1 + (1 - \alpha_1)e^{-\frac{x - \lambda}{\delta}}} \frac{\alpha_2 e^{-\frac{x - \lambda}{\delta}} e^{-e^{-\frac{x - \lambda}{\delta}}\delta}}{\delta \left( \alpha_2 + (1 - \alpha_2)e^{-\frac{x - \lambda}{\delta}} \right)^2} dx
\]

\[
= \frac{\alpha_1}{(\alpha_1 \alpha_2 - 1)^2} \left[ -\ln \frac{\alpha_1}{\alpha_2} + \frac{\alpha_1}{\alpha_2} - 1 \right]
\]

Let \((x_1, \ldots, x_m)\) and \((y_1, \ldots, y_n)\) be two independent random samples of sizes \( m \) and \( n \) from Marshall-Olkin Gumbel maximum distribution with Marshall-Olkin parameters \( \alpha_1 \) and \( \alpha_2 \), respectively, and common unknown parameters \( \lambda \) and \( \delta \). L is the log likelihood function, then maximum likelihood estimates of the unknown parameters \( \alpha_1, \alpha_2 \) are the solutions of the non-linear equations \( \frac{\partial L}{\partial \alpha_1} = 0 \) and \( \frac{\partial L}{\partial \alpha_2} = 0 \) respectively. The elements of information matrix are

\[
I_{11} = -E \left( \frac{\partial^2 L}{\partial \alpha_1^2} \right) = \frac{m}{3\alpha_1^2}
\]

Similarly,

\[
I_{22} = -E \left( \frac{\partial^2 L}{\partial \alpha_2^2} \right) = \frac{n}{3\alpha_2^2}
\]

\[
I_{12} = I_{21} = -E \left( \frac{\partial^2 L}{\partial \alpha_1 \partial \alpha_2} \right) = 0.
\]
By the property of m.l.e for $m \to \infty, n \to \infty$, we obtain that

$$(\sqrt{m}(\hat{\alpha}_1 - \alpha_1), \sqrt{n}(\hat{\alpha}_2 - \alpha_2))^T \overset{d}{\to} N_2\left(\mathbf{0}, \text{diag}\{a_{11}^{-1}, a_{22}^{-1}\}\right),$$

where $a_{11} = \lim_{m,n \to \infty} \frac{1}{m} I_{11} = \frac{1}{3\alpha_1^2}$ and $a_{22} = \lim_{m,n \to \infty} \frac{1}{n} I_{22} = \frac{1}{3\alpha_2^2}$. The 95% confidence interval for $R$ is given by

$$\hat{R} \pm 1.96 \hat{\alpha}_1 b_1(\hat{\alpha}_1, \hat{\alpha}_2) \sqrt{\frac{3}{m} + \frac{3}{n}},$$

where $\hat{R} = R(\hat{\alpha}_1, \hat{\alpha}_2)$ is the estimator of $R$ and

$$b_1(\alpha_1, \alpha_2) = \frac{\partial R}{\partial \alpha_1} = \frac{\alpha_2}{(\alpha_1 - \alpha_2)^3} \left[2(\alpha_1 - \alpha_2) + (\alpha_1 + \alpha_2) \log \frac{\alpha_2}{\alpha_1}\right].$$

### 2.7.1 Simulation Study

We generate $N = 1000$ sets of $X$ samples and $Y$ samples from Marshall-Olkin Gumbel maximum distribution with parameters $\alpha_1, \lambda, \delta$ and $\alpha_2, \lambda, \delta$ respectively. The combinations of samples of sizes $m = 20, 30, 40$ and $n = 20, 30, 40$ are considered. The estimates of $\alpha_1$ and $\alpha_2$ are then obtained from each sample to obtain $\hat{R}$. The validity of the estimate of $R$ is discussed by the measures:

1) Average bias of the simulated $N$ estimates of $R$:

$$\frac{1}{N} \sum_{i=1}^{N} (\hat{R}_i - R)$$

2) Average mean square error of the simulated $N$ estimates of $R$:

$$\frac{1}{N} \sum_{i=1}^{N} (\hat{R}_i - R)^2$$
Table 2.2: Average bias and average mean square error of the simulated estimates of \( R \) for \( \delta = 2.5, \lambda = 2 \)

<table>
<thead>
<tr>
<th>(m,n)</th>
<th>(1.5,1.4)</th>
<th>(0.5,1.5)</th>
<th>(2.0,5)</th>
<th>(0.8,0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(20,20)</td>
<td>-0.0082</td>
<td>0.1159</td>
<td>-0.1490</td>
<td>0.0159</td>
</tr>
<tr>
<td>(20,30)</td>
<td>-0.0168</td>
<td>0.1113</td>
<td>-0.1529</td>
<td>0.0089</td>
</tr>
<tr>
<td>(20,40)</td>
<td>-0.0165</td>
<td>0.1093</td>
<td>-0.1563</td>
<td>0.0037</td>
</tr>
<tr>
<td>(40,20)</td>
<td>-0.0037</td>
<td>0.1238</td>
<td>-0.1438</td>
<td>0.0220</td>
</tr>
</tbody>
</table>

Table 2.3: Average confidence length and coverage probability of the simulated estimates of \( R \) for \( \delta = 2.5, \lambda = 2 \)

<table>
<thead>
<tr>
<th>(m,n)</th>
<th>(1.5,1.4)</th>
<th>(0.5,1.5)</th>
<th>(2.0,5)</th>
<th>(0.8,0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(20,20)</td>
<td>0.3546</td>
<td>0.3490</td>
<td>0.3478</td>
<td>0.3545</td>
</tr>
<tr>
<td>(20,30)</td>
<td>0.3242</td>
<td>0.3484</td>
<td>0.3165</td>
<td>0.3241</td>
</tr>
<tr>
<td>(20,40)</td>
<td>0.3245</td>
<td>0.3185</td>
<td>0.3194</td>
<td>0.3240</td>
</tr>
<tr>
<td>(40,20)</td>
<td>0.3078</td>
<td>0.3019</td>
<td>0.3035</td>
<td>0.3075</td>
</tr>
</tbody>
</table>

3) Average length of the asymptotic 95% confidence intervals of \( R \):

\[
\frac{1}{N} \sum_{i=1}^{N} 2(1.96) b_i \left( \hat{\alpha}_{1i}, \hat{\alpha}_{2i} \right) \sqrt{ \frac{3}{m} + \frac{3}{n} }
\]

4) The coverage probability of the \( N \) simulated confidence intervals given by the proportion of such interval that include the parameter \( R \).

Average bias and average mean square error of the simulated estimates of \( R \) for \( \delta = 2.5, \lambda = 2 \) is given in Table 2.2 and average confidence length and coverage probability of the simulated estimates of \( R \) for \( \delta = 2.5, \lambda = 2 \) is given in Table 2.3.
Table 2.4: Estimated values, loglikelihood, Kolmogrove-Smirnov statistic and P value for the data set are given

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Estimates</th>
<th>$-\log L$</th>
<th>$K - S$</th>
<th>$P - value$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumbel</td>
<td>$\hat{\lambda} = 2155.1632$</td>
<td>886.5109</td>
<td>0.0763</td>
<td>0.5374</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta} = 635.5164$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshall-Olkin Gumbel</td>
<td>$\hat{\lambda} = 1.6106$</td>
<td>885.0302</td>
<td>0.0572</td>
<td>0.8615</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta} = 398.2324$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha} = 485.094$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.8 Data Analysis and Modeling

In this section we analyze a real data set and compare Marshall Olkin Gumbel maximum distribution with Gumbel maximum distribution. The data gives the annual maximum daily discharge of Rhone River (1826-1936), Gumbel (1941). The P-P and Q-Q plots for the two distributions are given in Fig 2.2 and in Fig 2.3. Estimated values are given in table 2.4. From that we can conclude that the Marshall Olkin Gumbel maximum distribution is a better fit. Estimated values, loglikelihood, Kolmogrove-Smirnov statistic and P value for the data set are given in Table 2.4. Simulated sample path of the Marshall-Olkin Gumbel minification processes for various values of the parameters are given in Fig 2.4a, b, c, d. We derive the first order autocorrelations of the Marshall-Olkin Gumbel minification processes using Monte Carlo method. For different values of the parameter $\lambda$, $\delta$ and $\alpha$ we simulate 100 sequences of 1000 observations from the process. For each sequence we estimate the sample first order autocorrelation. In Table 2.5 we illustrate the averages of these sample first order autocorrelations with standard deviations in brackets.

2.9 Conclusions

In this chapter Marshall-Olkin Gumbel distribution is introduced and its properties are studied. Minification processes with Marshall-Olkin Gumbel marginal distribution are also developed and studied. We analyze a real data set and compare Marshall Olkin Gumbel
**Figure 2.2:** The P-P and Q-Q plots of Gumbel maximum distribution

**Figure 2.3:** The P-P and Q-Q plots of Marshall-Olkin Gumbel maximum distribution
Figure 2.4: Simulated sample path of the Marshall-Olkin Gumbel minification processes for various values of the parameters

Fig 2.4a. when \( n = 1000, \lambda = 20, \delta = 5, \alpha = .9 \).

Fig 2.4b. when \( n = 500, \lambda = 2, \delta = 10, \alpha = 5 \).

Fig 2.4c. when \( n = 300, \lambda = 30, \delta = 10, \alpha = 10 \).

Fig 2.4d. when \( n = 500, \lambda = 2, \delta = 5, \alpha = .5 \).
Table 2.5: The sample first order autocorrelations of the Marshall-Olkin Gumbel mini-
fication process for different values of $\lambda$ and $\alpha$ when $\delta=0.9$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>0.01</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.81599</td>
<td>0.65497</td>
<td>0.55365</td>
<td>0.43169</td>
<td>0.342256</td>
<td>0.27767</td>
<td>0.217256</td>
<td>0.16591</td>
</tr>
<tr>
<td></td>
<td>(0.06654)</td>
<td>(0.05599)</td>
<td>(0.04895)</td>
<td>(0.04469)</td>
<td>(0.03729)</td>
<td>(0.04594)</td>
<td>(0.03658)</td>
<td>(0.03594)</td>
</tr>
<tr>
<td>4</td>
<td>0.81407</td>
<td>0.66666</td>
<td>0.55866</td>
<td>0.42796</td>
<td>0.35589</td>
<td>0.27512</td>
<td>0.22255</td>
<td>0.16702</td>
</tr>
<tr>
<td></td>
<td>(0.06690)</td>
<td>(0.04935)</td>
<td>(0.04106)</td>
<td>(0.04818)</td>
<td>(0.04197)</td>
<td>(0.04040)</td>
<td>(0.03429)</td>
<td>(0.035321)</td>
</tr>
<tr>
<td>6</td>
<td>0.80175</td>
<td>0.65714</td>
<td>0.56561</td>
<td>0.43875</td>
<td>0.34779</td>
<td>0.27634</td>
<td>0.21097</td>
<td>0.16438</td>
</tr>
<tr>
<td></td>
<td>(0.08137)</td>
<td>(0.05982)</td>
<td>(0.04215)</td>
<td>(0.042311)</td>
<td>(0.03458)</td>
<td>(0.03549)</td>
<td>(0.03635)</td>
<td>(0.03653)</td>
</tr>
<tr>
<td>8</td>
<td>0.81002</td>
<td>0.65178</td>
<td>0.55423</td>
<td>0.43687</td>
<td>0.34622</td>
<td>0.27415</td>
<td>0.208935</td>
<td>0.170307</td>
</tr>
<tr>
<td></td>
<td>(0.07088)</td>
<td>(0.05569)</td>
<td>(0.04782)</td>
<td>(0.04364)</td>
<td>(0.03533)</td>
<td>(0.03763)</td>
<td>(0.03888)</td>
<td>(0.03832)</td>
</tr>
<tr>
<td>10</td>
<td>0.81526</td>
<td>0.65276</td>
<td>0.56661</td>
<td>0.43452</td>
<td>0.34819</td>
<td>0.27226</td>
<td>0.214252</td>
<td>0.15743</td>
</tr>
<tr>
<td></td>
<td>(0.06641)</td>
<td>(0.05211)</td>
<td>(0.04483)</td>
<td>(0.05002)</td>
<td>(0.03291)</td>
<td>(0.03977)</td>
<td>(0.03333)</td>
<td>(0.03344)</td>
</tr>
<tr>
<td>15</td>
<td>0.82041</td>
<td>0.66538</td>
<td>0.55676</td>
<td>0.43982</td>
<td>0.34859</td>
<td>0.28724</td>
<td>0.218361</td>
<td>0.16336</td>
</tr>
<tr>
<td></td>
<td>(0.07283)</td>
<td>(0.05726)</td>
<td>(0.05979)</td>
<td>(0.04446)</td>
<td>(0.03549)</td>
<td>(0.03281)</td>
<td>(0.038377)</td>
<td>(0.03408)</td>
</tr>
<tr>
<td>20</td>
<td>0.81810</td>
<td>0.64376</td>
<td>0.56149</td>
<td>0.43054</td>
<td>0.346518</td>
<td>0.27348</td>
<td>0.217007</td>
<td>0.16152</td>
</tr>
<tr>
<td></td>
<td>(0.06479)</td>
<td>(0.06172)</td>
<td>(0.04725)</td>
<td>(0.04285)</td>
<td>(0.039740)</td>
<td>(0.03632)</td>
<td>(0.035260)</td>
<td>(0.031374)</td>
</tr>
<tr>
<td>25</td>
<td>0.80386</td>
<td>0.65189</td>
<td>0.568634</td>
<td>0.434051</td>
<td>0.349032</td>
<td>0.276988</td>
<td>0.217397</td>
<td>0.16749</td>
</tr>
<tr>
<td></td>
<td>(0.08072)</td>
<td>(0.05511)</td>
<td>(0.049378)</td>
<td>(0.04668)</td>
<td>(0.040416)</td>
<td>(0.035309)</td>
<td>(0.03478)</td>
<td>(0.03558)</td>
</tr>
<tr>
<td>30</td>
<td>0.82566</td>
<td>0.65373</td>
<td>0.55509</td>
<td>0.42861</td>
<td>0.346023</td>
<td>0.27259</td>
<td>0.220558</td>
<td>0.164879</td>
</tr>
<tr>
<td></td>
<td>(0.06395)</td>
<td>(0.05378)</td>
<td>(0.04579)</td>
<td>(0.04102)</td>
<td>(0.03799)</td>
<td>(0.03384)</td>
<td>(0.03557)</td>
<td>(0.03241)</td>
</tr>
</tbody>
</table>

maximum distribution with Gumbel maximum distribution. We conclude that Marshall Olkin Gumbel maximum distribution is a better fit. Estimation of reliability is also done. A minification process is constructed and its covariance structure is derived. The processes developed can be applied for modeling data on climate changes such as maximum temperature, minimum temperature, wind velocity, rainfall data, humidity levels as well as environmental statistics.

References


