Ann. Inst. Statist. Math.
Vol. 51, No. 2, 253–264 (1999)

MOMENT ESTIMATION FOR MULTIVARIATE EXTREME VALUE DISTRIBUTION IN A NESTED LOGISTIC MODEL*

DAOJI SHI[1] AND SHENGSHENG ZHOU[2]

[1]*Department of Mathematics, Tianjin University, Tianjin 300072, China*
[2]*Department of Mathematics, Anhui Institute of Mechanical and Electrical Engineering, Wuhu 241000, China*

(Received November 1, 1995; revised February 23, 1998)

Abstract. This paper considers multivariate extreme value distribution in a nested logistic model. The dependence structure for this model is discussed. We find a useful transformation that transformed variables possess the mixed independence. Thus, the explicit algebraic formulae for a characteristic function and moments may be given. We use the method of moments to derive estimators of the dependence parameters and investigate the properties of these estimators in large samples via asymptotic theory and in finite samples via computer simulation. We also compare moment estimation with a maximum likelihood estimation in finite sample sizes. The results indicate that moment estimation is good for all practical purposes.

Key words and phrases: Gumbel distribution, maximum likelihood estimation, moment estimation, multivariate extreme value distribution.

1. Introduction

Statistical methods for multivariate extremes have been used to solve many engineering problems (for example, see Coles and Tawn (1994)), but parametric estimation for p-variate ($p \geq 3$) extreme value distribution has received only limited attention. Because no natural parametric family exists for the dependence structure of the multivariate extreme value distribution, this must be modelled in some way. A number of parametric models have been given by Coles and Tawn (1991). The logistic model is one of the essential parametric models which has the joint distribution function

$$G(x_1, \ldots, x_p) = \exp\left\{-\left(\sum_{i=1}^{p} e^{-x_i/\alpha}\right)^{\alpha}\right\}, \tag{1.1}$$

* The work was supported by the National Natural Science Foundation of China.

where we assume, without loss of generality, that the margins have standard Gumbel distribution

$$H(x) = \exp(-e^{-x}).$$

Some papers have been written concerning parametric estimation in this model, for example, for the asymptotic behaviour of a maximum likelihood estimation, see Tawn (1988), Oakes and Manatunga (1992), Shi *et al.* (1992), and Shi (1995*b*, 1995*c*), while a moment estimation has been presented by Shi (1995*a*).

It requires a single parameter, α, to govern dependence among the variates in a logistic model. Such a model would not be very realistic in many cases. McFadden (1978) and Tawn (1990) developed a class of models which involves hierarchical dependence, i.e. a nested logistic model. In order to simplify the expressions, here we consider only trivariate extreme value distribution. This is given by the joint distribution function

$$F(x_1, x_2, x_3) = \exp[-\{(e^{-x_1/\alpha\beta} + e^{-x_2/\alpha\beta})^\beta + e^{-x_3/\alpha}\}^\alpha], \tag{1.2}$$

where $0 \le \alpha, \beta \le 1$ are dependence parameters. It is clear that (1.1) is a special case of (1.2) with $\beta = 1$. When $\alpha = 1$, (1.2) may be reduced to the bivariate extreme value distribution in the logistic model. Tawn (1990) described a physical motivation for the nested logistic model.

To our knowledge, no results have yet been published within the context of parametric estimation of the nested logistic model. Here we consider the moment estimation of the dependence parameters in the model (1.2). Moment estimation has some nice properties, for example, it is computationally simple, consistent and asymptotically normal. In Section 2, we analyse the dependence structure of the model and give a useful transformation so that we can derive an explicit algebraic expression of the moment estimator in Section 3, and an approximately asymptotic covariance matrix in Section 4. We use computer simulation to obtain the biases and the variance-covariance of the estimation and compare these with the maximum likelihood estimation in finite sample sizes in Section 5.

2. Dependence structure of model

For the sake of brevity, let

$$s_i = e^{-x_i}, \qquad i = 1, 2, 3,$$

so that (1.2) is reduced as

$$F(x_1, x_2, x_3) = \exp[-\{(s_1^{1/\alpha\beta} + s_2^{1/\alpha\beta})^\beta + s_3^{1/\alpha}\}^\alpha],$$

and the density of the distribution (1.2) may be written as

$$\begin{aligned} f(x_1, x_2, x_3) &= \frac{\partial F^3(x_1, x_2, x_3)}{\partial x_1 \partial x_2 \partial x_3} \\ &= \frac{1}{\sigma_1\sigma_2\sigma_3} e^{-u} u^{1-2/\alpha} v^{1/\alpha - 2/\alpha\beta} (s_1 s_2)^{1/\alpha\beta} s_3^{1/\alpha} Q, \end{aligned} \tag{2.1}$$

where

$$
\begin{aligned}
&v = (s_1^{1/\alpha\beta} + s_2^{1/\alpha\beta})^{\alpha\beta},\\
&u = \{(s_1^{1/\alpha\beta} + s_2^{1/\alpha\beta})^{\beta} + s_3^{1/\alpha}\}^{\alpha} = (v^{1/\alpha} + s_3^{1/\alpha})^{\alpha},\\
&Q = \left(\frac{v}{u}\right)^{1/\alpha} Q_3(u;\alpha) + \frac{1-\beta}{\alpha\beta} Q_2(u;\alpha),
\end{aligned}
$$

$$Q_2(u;\alpha) = u + \frac{1}{\alpha} - 1, \tag{2.2}$$

$$Q_3(u;\alpha) = u^2 + 3\left(\frac{1}{\alpha} - 1\right) u + \left(\frac{1}{\alpha} - 1\right)\left(\frac{2}{\alpha} - 1\right). \tag{2.3}$$

Especially when $\beta = 1$, i.e. in the logistic model, we have the density of the distribution (1.1)

$$g(x_1, x_2, x_3) = \frac{\partial G^3(x_1, x_2, x_3)}{\partial x_1 \partial x_2 \partial x_3} = \frac{(s_1 s_2 s_3)^{1/\alpha}}{\sigma_1 \sigma_2 \sigma_3} e^{-u} u^{1-3/\alpha} Q_3(u;\alpha).$$

In general, p-variate extreme value distribution in the logistic model has the density (Shi (1995c))

$$g(x_1, \dots, x_p) = \left(\prod_{i=1}^{p} \frac{s_i^{1/\alpha}}{\sigma_i}\right) u^{1-p/\alpha} Q_p(u;\alpha) e^{-u},$$

where

$$
\begin{aligned}
&Q_p(u;\alpha) = \left(\frac{p-1}{\alpha} + u - 1\right) Q_{p-1}(u;\alpha) - u\frac{\partial Q_{p-1}(u;\alpha)}{\partial u},\\
&Q_1(u;\alpha) = 1.
\end{aligned}
$$

Make a transformation

$$
\begin{cases}
s_1 = u t_1^{\alpha} t_2^{\alpha\beta},\\
s_2 = u t_1^{\alpha} (1 - t_2)^{\alpha\beta},\\
s_3 = u (1 - t_1)^{\alpha},
\end{cases}
$$

or

$$
\begin{cases}
x_1 = -\log(u t_1^{\alpha} t_2^{\alpha\beta}),\\
x_2 = -\log\{u t_1^{\alpha} (1 - t_2)^{\alpha\beta}\},\\
x_3 = -\log\{u (1 - t_1)^{\alpha}\}, \qquad u > 0, \quad 0 < t_1, t_2 < 1.
\end{cases} \tag{2.4}
$$

We can regard u as a 'distance' from the origin to (s_1, s_2, s_3) and t_1, t_2 as a dependence between s_1, s_2, s_3. We find that the Jacobian of transformation (2.4) is

$$\frac{\partial(x_1, x_2, x_3)}{\partial(u, t_1, t_2)} = \frac{\sigma_1 \sigma_2 \sigma_3 \alpha^2 \beta}{u t_1 t_2 (1 - t_1)(1 - t_2)}.$$

Therefore, the joint density of the transformed random vector (U, T_1, T_2) is

$$w(u, t_1, t_2) = \alpha^2 \beta Q e^{-u} = \beta r_3(u, t_1; \alpha) + (1-\beta) r_2(u, t_1; \alpha), \tag{2.5}$$

where, for $p = 2, 3$,

$$r_p(u, t_1; \alpha) = h_p(u; \alpha) a_p(t_1)$$

and

$$h_p(u; \alpha) = \frac{\alpha^{p-1}}{(p-1)!} Q_p(u; \alpha) e^{-u} = \sum_{j=1}^{p} q_{pj} \gamma(u, j)$$

is the density of mixed gamma distribution and $\gamma(u, j)$ denotes the gamma distribution density with parameter j, $\gamma(u, j) = u^{j-1} e^{-u} / \Gamma(j)$, $u > 0$. The weighting coefficient q_{pj} has been given by Shi (1995c). For example, we have $q_{21} = 1 - \alpha$, $q_{22} = \alpha$; $q_{31} = \frac{1}{2}(1-\alpha)(2-\alpha)$, $q_{32} = \frac{3}{2}\alpha(1-\alpha)$, $q_{33} = \alpha^2$, and

$$a_p(t) = (p-1) t^{p-2}, \qquad 0 < t < 1$$

is the density of the power function distribution on interval $(0, 1)$.

From (2.5), it can be seen that T_2 is independent of (U, T_1). T_2 has the uniform distribution on interval $(0, 1)$, and (U, T_1) is the $\beta : 1 - \beta$ mixture of $r_3(u, t_1; \alpha)$ and $r_2(u, t_1; \alpha)$. In any $r_p(u, t_1; \alpha)$, $p = 2, 3$, U is independent of T_1. We call this relation a mixed independence. So, in the nested logistic model the transformed variables U, T_1, and T_2 possess mixed independence.

Mixed independence shows that the dependence between s_i and s_j (therefore x_i and x_j) is absolutely determined by independent variates T_1 and T_2. Under mixed independence, it is easy to find the expectation of a function with the form $G_1(U)G_2(T_1)G_3(T_2)$. This is obtained directly:

$$\begin{aligned} &E[G_1(U)G_2(T_1)G_3(T_2)] \\ &= \int_0^1 \int_0^\infty G_1(u)G_2(t_1) \\ &\qquad \cdot [\beta h_3(u; \alpha) g_3(t_1) + (1-\beta) h_2(u; \alpha) g_2(t_1)] du dt_1 \\ &\quad \cdot \int_0^1 G_3(t_2) dt_2 \\ &= \{\beta E_3[G_1(U)] E_3[G_2(T_1)] \\ &\quad + (1-\beta) E_2[G_1(U)] E_2[G_2(T_1)]\} E[G_3(T_2)], \end{aligned} \tag{2.6}$$

where

$$E_p[G_1(U)] = \sum_{j=1}^{p} \frac{q_{pj}}{\Gamma(j)} \int_0^\infty G_1(u) u^{j-1} e^{-u} du, \qquad p = 2, 3, \tag{2.7}$$

$$E_p[G_2(T_1)] = \int_0^1 (p-1) G_2(t_1) t_1^{p-2} dt_1, \qquad p = 2, 3, \tag{2.8}$$

$$E[G_3(T_2)] = \int_0^1 G_3(t_2) dt_2. \tag{2.9}$$

It should be apparent from (2.7), (2.8) and (2.9) that $E_p[G_1(U)]$, $E_p[G_2(T_1)]$ and $E[G_3(T_2)]$ are easier to calculate and may be expressed in explicit formulae for most practical cases. Therefore, when G_1, G_2, G_3 are some simple functions the expectation of their product may have explicit expression through (2.6). For example, we have

$$E[U^a T_1^b (1-T_1)^c] = (1+a\alpha)\Gamma(1+a)\Gamma(1+b)\Gamma(1+c) \cdot \frac{\beta(2+a\alpha)\frac{1+b}{2+b+c}+1-\beta}{\Gamma(2+b+c)}. \tag{2.10}$$

Especially when $b=c=0$, we have

$$E(U^a) = (1+a\alpha)\left(1+\frac{a\alpha\beta}{2}\right)\Gamma(1+a).$$

Differentiating with respect to a in both side of the above expression, $E(U^a \log U)$ may be obtained. Similarly, the expectations evolving in $\log T_1$, $\log(1-T_1)$ may also be calculated by differentiating with respect to b and c in both side of (2.10), respectively.

3. Characteristic function and moment

Using the mixed independence of the nested logistic model, we can derive a characteristic function and moments in an explicit expression. The characteristic function of the distribution (1.2) follows from (2.10):

$$\begin{aligned}\Phi(t_1,t_2,t_3) = E\{e^{i(t_1X_1+t_2X_2+t_3X_3)}\} &= e^{it'\mu}\frac{\Gamma(1-it'\sigma)}{\Gamma(1-i\alpha t'\sigma)} \\ &\times \frac{\Gamma[1-i\alpha(t_1\sigma_1+t_2\sigma_2)]}{\Gamma[1-i\alpha\beta(t_1\sigma_1+t_2\sigma_2)]} \\ &\times \Gamma(1-i\alpha\beta t_1\sigma_1)\Gamma(1-i\alpha\beta t_2\sigma_2)\Gamma(1-i\alpha t_3\sigma_3),\end{aligned} \tag{3.1}$$

here $t' = (t_1,t_2,t_3)$, $\mu' = (\mu_1,\mu_2,\mu_3)$, $\sigma' = (\sigma_1,\sigma_2,\sigma_3)$. When $\beta = 1$, (3.1) reduces the characteristic function in the logistic model (see Shi (1995*a*)).

From (3.1), we easily find that

$$\left.\frac{\partial\Phi}{\partial t_j}\right|_{t=0} = i(\mu_j + \sigma_j\gamma), \qquad j=1,2,3,$$

where i is the imaginary unit and γ is Euler's constant. The relation between the characteristic function and moment yields

$$E(X_j) = \mu_j + \sigma_j\gamma, \qquad j=1,2,3.$$

On the other hand, we can also directly calculate any moments of the distribution (1.2). For example, from (2.4) we have

$$E(X_1) = -E(\log U + \alpha\log T_1 + \alpha\beta\log T_2).$$

It is necessary to calculate $E(\log U)$, $E(\log T_1)$, $E(\log T_2)$. From (2.6) we have

$$\begin{aligned} E(\log U) &= \beta E_3(\log U) + (1-\beta)E_2(\log U) \\ &= \beta\left(\frac{3\alpha}{2} - \gamma\right) + (1-\beta)(\alpha-\gamma) = \frac{\alpha\beta}{2} + \alpha - \gamma, \end{aligned}$$

where $E_3(\log U) = \frac{3\alpha}{2} - \gamma$, $E_2(\log U) = \alpha - \gamma$ can be obtained from (2.7). Similarly, from (2.8) and (2.6), we have $E(\log T_1) = \frac{\beta}{2} - 1$. And $E(\log T_2) = -1$ is directly obtained from (2.9) and (2.6). It is immediately obvious from the results given above that $E(X_1) = \mu_1 + \sigma_1\gamma$.

More calculation may derive any moment of the distribution (1.2). For example, we have

$$\begin{aligned} &E[X_j - E(X_j)]^2 = \sigma_j^2\frac{\pi^2}{6}, \quad j = 1,2,3, \\ &E[X_1 - E(X_1)][X_2 - E(X_2)] = \sigma_1\sigma_2\frac{\pi^2}{6}(1-\alpha^2\beta^2), \\ &E[X_j - E(X_j)][X_3 - E(X_3)] = \sigma_j\sigma_3\frac{\pi^2}{6}(1-\alpha^2), \quad j = 1,2 \end{aligned}$$

(also see Tiago de Oliveira (1980)). Therefore, the correlation coeffcient must be

$$\begin{aligned} \rho_{12} &= \mathrm{Corr}(X_1, X_2) = 1 - \alpha^2\beta^2, \\ \rho_{j3} &= \mathrm{Corr}(X_j, X_3) = 1 - \alpha^2, \quad j = 1,2. \end{aligned}$$

The linear dependence degree between x_i and x_j is completely determined by the dependence parameters α, β. Now we may obtain the moment estimators of the parameters in the distribution (1.2):

$$\hat{\alpha} = \frac{1}{2}(\sqrt{1-r_{13}} + \sqrt{1-r_{23}}), \tag{3.2}$$

$$\hat{\beta} = \frac{\sqrt{1-r_{12}}}{\hat{\alpha}}, \tag{3.3}$$

where r_{ij}, $i < j$, $i,j = 1,2,3$ are the usual sample correlations. Of course, we must ensure that $0 \le \hat{\alpha}, \hat{\beta} \le 1$.

It needs higher moments of the distribution (1.2) to give a covariance matrix of the moment estimators $\hat{\alpha}$, $\hat{\beta}$. We see that X_1, X_2 are symmetric in the distribution (1.2) and X_3 is different from X_1, X_2. Consequently, the moments concerned in X_3 have different expressions. Let

$$\mu_{ijk} = E[(X_1 - EX_1)^i (X_2 - EX_2)^j (X_3 - EX_3)^k].$$

We have the following results about the third moment

$$\mu_{ijk} = 2\sigma_1^i\sigma_2^j\sigma_3^k(1-\lambda^3)\zeta_3, \quad i+j+k = 3, \quad 0 \le i,j,k \le 3,$$

where $\lambda = 0$ in the case where there are two zeros among i, j, k; or else there is at most one zero among i, j, k, $\lambda = \alpha$ if $k \neq 0$; $\lambda = \alpha\beta$ if $k = 0$. And $\zeta_3 = 1.20205690\cdots$ is the zeta function, $\zeta_s = \sum_{k=1}^{\infty} k^{-s}$, evaluated at $s = 3$ (see Abramowitz and Stegun (1964, §23.2)).

Similarly, for fourth moments, i.e. $i + j + k = 4$, we have

$$\mu_{ijk} = \frac{1}{60}\sigma_1^i\sigma_2^j\sigma_3^k(9 + 4\lambda^2)(1 - \lambda^2)\pi^4$$

when either i, j, and k are not equal or there are two zeros among i, j, k. In the other case, there is equality among i, j and k, and we have

$$\mu_{ijk} = \sigma_1^i\sigma_2^j\sigma_3^k\left\{\frac{1 - \delta_{kl}^2}{36} + \frac{(11 - 5\eta_{kl}^2 + 6\kappa_{kl}^2)(1 - \kappa_{kl}^2)}{90}\right\}\pi^4,$$

where $l = 1$, if $i \neq j$; $l = 2$, if $i = j$, and

$$\begin{array}{lll} \delta_{11} = \alpha, & \eta_{11} = \alpha\beta, & \kappa_{11} = \alpha, \\ \delta_{21} = 0, & \eta_{21} = \alpha, & \kappa_{21} = \alpha, \\ \delta_{22} = \alpha\beta, & \eta_{22} = \alpha, & \kappa_{22} = \alpha, \\ \delta_{02} = 0, & \eta_{02} = \alpha\beta, & \kappa_{02} = \alpha\beta. \end{array}$$

4. Asymptotic covariance matrix of moment estimation

In general, moment estimators are simple and often useful in applications. Moment estimation possesses consistent asymptotic normality. Except in some special cases, it is impossible to accurately find the covariance between the estimators. The method described by Kendall and Stuart ((1969), §10.6) can be used to obtain approximately the asymptotic covariance. Using the expressions for moments given in the preceding section, for example, we have (approximate to $o(n^{-1})$)

$$\operatorname{cov}(S_1, S_2) = \frac{\mu_{220} - \mu_{200}\mu_{020}}{4n\sqrt{\mu_{200}\mu_{020}}} = \frac{1}{60n}\sigma_1\sigma_2(11 + \alpha^2\beta^2)(1 - \alpha^2\beta^2)\pi^2,$$

but

$$\operatorname{cov}(S_1, S_3) = \frac{\mu_{202} - \mu_{200}\mu_{002}}{4n\sqrt{\mu_{200}\mu_{002}}} = \frac{1}{60n}\sigma_1\sigma_3(11 + \alpha^2)(1 - \alpha^2)\pi^2.$$

In an analogous way, we have

$$\begin{aligned} \operatorname{cov}(r_{12}, r_{13}) = \frac{\rho_{12}\rho_{13}}{n}&\left\{\frac{1}{4}\left(\frac{\mu_{400}}{\mu_{200}^2} + \frac{\mu_{202}}{\mu_{200}\mu_{002}} + \frac{\mu_{220}}{\mu_{200}\mu_{020}} + \frac{\mu_{022}}{\mu_{020}\mu_{002}}\right)\right. \\ &- \frac{1}{2}\left(\frac{\mu_{301}}{\mu_{200}\mu_{101}} + \frac{\mu_{121}}{\mu_{020}\mu_{101}} + \frac{\mu_{310}}{\mu_{200}\mu_{110}} + \frac{\mu_{112}}{\mu_{002}\mu_{110}}\right) \\ &\left. + \frac{\mu_{211}}{\mu_{110}\mu_{101}}\right\} \\ = \frac{1 - \alpha^2}{10n}&\alpha^4\beta^2(11\beta^2 + 44 + 2\alpha^2 + \alpha^2\beta^4), \end{aligned}$$

but

$$\operatorname{cov}(r_{13}, r_{23}) = \frac{1-\alpha^2}{10n}\alpha^4(55+3\alpha^2),$$

and so on. After tedious calculation, it turns out that approximate expressions for elements of the asymptotic covariance matrix of the estimators of the dependence parameters given in (3.2) and (3.3) (approximate to $o(n^{-1})$) are

$$D(\hat{\alpha}) = \frac{\alpha^2}{80n}(131 - 116\alpha^2 - 5\alpha^4), \tag{4.1}$$

$$D(\hat{\beta}) = \frac{\beta^2}{80n}(107 + 52\alpha^2 - 44\beta^2 + 3\alpha^4 - 84\alpha^2\beta^2 - 4\alpha^2\beta^4), \tag{4.2}$$

$$\operatorname{cov}(\hat{\alpha}, \hat{\beta}) = \frac{\alpha\beta}{80n}(-43 + 32\alpha^2 + 22\beta^2 - 22\alpha^2\beta^2 + \alpha^4 + 2\alpha^2\beta^4 - 2\alpha^4\beta^4).$$

5. Comparison with maximum likelihood estimation

In this section, we want to investigate the properties of the moment estimation in finite samples and to compare them with maximum likelihood estimation by simulation. So far we have little knowledge about maximum likelihood estimation for the parameters of the distribution (1.2), even its asymptotic behaviour, because of its complicated form. Now we consider maximum likelihood estimation for dependence parameters. Firstly, we can write the log-likelihood function, denoting by ℓ, for a single observation from (2.1)

$$\begin{aligned} \ell &= \log f(x_1, x_2, x_3) \\ &= -\log(\sigma_1\sigma_2\sigma_3) - u + \left(1 - \frac{2}{\alpha}\right)\log u + \left(\frac{1}{\alpha} - \frac{2}{\alpha\beta}\right)\log v \\ &\quad + \frac{1}{\alpha\beta}\log(s_1 s_2) + \frac{1}{\alpha}\log s_3 + \log Q. \end{aligned} \tag{5.1}$$

The score statistic may be expressed as

$$\begin{aligned} \frac{\partial \ell}{\partial \alpha} &= -\frac{1}{\alpha}\log[t_1 t_2(1-t_1)(1-t_2)] + \left(u + \frac{2}{\alpha} - 1\right)D - \frac{\beta - 2}{\alpha}C_2 \\ &\quad + \frac{1}{Q}\left\{\frac{Q_3}{\alpha}(Dt_1 - \beta t_1 C_2 - t_1\log t_1) - uDt_1\frac{\partial Q_3}{\partial u}\right. \\ &\qquad \left. - \frac{1-\beta}{\alpha\beta}uD + t_1\frac{\partial Q_3}{\partial \alpha} - \frac{1-\beta}{\alpha^2\beta}Q_2 + \frac{1-\beta}{\alpha\beta}\frac{\partial Q_2}{\partial \alpha}\right\}, \\ \frac{\partial \ell}{\partial \beta} &= (\alpha u - \alpha + 2)t_1 C_2 - \left(1 - \frac{2}{\beta}\right)C_2 - \frac{1}{\beta}\log[t_2(1-t_2)] \\ &\quad + \frac{1}{Q}\left\{-t_1(1-t_1)C_2 Q_3 - \alpha u t_1^2 C_2\frac{\partial Q_3}{\partial u} + \left(1 - \frac{1}{\beta}\right)u t_1 C_2 - \frac{Q_2}{\alpha\beta^2}\right\}, \end{aligned}$$

where

$$\begin{aligned} &C(t) = t\log t + (1-t)\log(1-t), \\ &C_1 = C(t_1), \quad C_2 = C(t_2), \quad D = C_1 + \beta t_1 C_2, \end{aligned}$$

Table 1. Simulated bias.

β	α / n	0.2		0.5		0.8	
		$b(\hat{\alpha})$	$b(\hat{\beta})$	$b(\hat{\alpha})$	$b(\hat{\beta})$	$b(\hat{\alpha})$	$b(\hat{\beta})$
		$b(\tilde{\alpha})$	$b(\tilde{\beta})$	$b(\tilde{\alpha})$	$b(\tilde{\beta})$	$b(\tilde{\alpha})$	$b(\tilde{\beta})$
0.2	30	.003	.004	.007	.005	−.001	.006
		.000	.003	.001	.004	−.005	.005
	50	.002	.002	.005	.003	−.001	.004
		.000	.001	.001	.001	−.003	.003
	100	.001	.001	.003	.002	−.001	.002
		.000	.000	.000	.001	−.002	.002
	200	.000	.000	.000	.001	−.000	.001
		.000	.000	.000	.000	−.001	.001
0.5	30	.003	.012	.006	.012	−.001	.015
		.000	.008	.000	.007	−.002	.009
	50	.002	.007	.004	.006	−.001	.010
		.000	.004	.000	.002	−.001	.007
	100	.001	.004	.002	.003	−.001	.006
		.000	.002	.000	.000	−.001	.004
	200	.001	.002	.001	.001	−.001	.004
		.000	.002	.000	.000	−.001	.003
0.8	30	.003	.010	.006	.010	−.001	.012
		.001	.004	.002	.004	−.001	.005
	50	.002	.006	.004	.005	−.001	.006
		.000	.004	.001	.004	−.001	.004
	100	.001	.003	.002	.002	−.001	.003
		.000	.002	.001	.001	−.001	.003
	200	.001	.002	.001	.001	−.001	.001
		.000	.001	.001	.001	−.001	.002

and Q_2, Q_3, as in (2.2), (2.3).

Let ℓ_n denote the log-likelihood ℓ based on n observations and $\tilde{\alpha}$, $\tilde{\beta}$ denote the maximum likelihood estimators which have be calculated by maximizing ℓ_n.

A computer simulation experiment was run to study the properties of the moment estimation and compare them with the maximum likelihood estimation in finite sample sizes. The mixed independence suggests a simple and accurate procedure to generate a random vector from the multivariate extreme value distribution (1.2). For example, see Shi *et al.* (1997). Simulations were performed for sample sizes $n = 30, 50, 100, 200$ with the dependence parameters $\alpha, \beta = 0.2, 0.5, 0.8$. For each combination of values of α, β and n, 5000 random samples from the distribution (1.2) were genereted, and the moment estimators, $\hat{\alpha}$, $\hat{\beta}$ and maximum likelihood estimators $\tilde{\alpha}$, $\tilde{\beta}$ were obtained for each sample, respectively. Quasi-Newton procedures work well for maximizing the log likelihood. This is based on iterative procedures. Here we take moment estimates as starting values.

Table 2. Simulated covariance matrix ($n \times$ covariance).

β	α / n	0.2 $D(\hat\alpha)$ / $D(\tilde\alpha)$	$D(\hat\beta)$ / $D(\tilde\beta)$	$C(\hat\alpha,\hat\beta)$ / $C(\tilde\alpha,\tilde\beta)$	0.5 $D(\hat\alpha)$ / $D(\tilde\alpha)$	$D(\hat\beta)$ / $D(\tilde\beta)$	$C(\hat\alpha,\hat\beta)$ / $C(\tilde\alpha,\tilde\beta)$	0.8 $D(\hat\alpha)$ / $D(\tilde\alpha)$	$D(\hat\beta)$ / $D(\tilde\beta)$	$C(\hat\alpha,\hat\beta)$ / $C(\tilde\alpha,\tilde\beta)$
.2	30	.07	.07	−.03	.38	.08	−.07	.47	.08	−.07
		.03	.06	−.03	.17	.06	−.07	.34	.05	−.08
	50	.07	.06	−.03	.37	.07	−.06	.51	.08	−.08
		.03	.05	−.03	.16	.05	−.07	.36	.05	−.09
	100	.07	.06	−.03	.37	.07	−.07	.52	.08	−.07
		.03	.05	−.03	.16	.05	−.07	.35	.05	−.09
	200	.07	.06	−.03	.37	.07	−.05	.53	.07	−.07
		.03	.05	−.03	.16	.05	−.06	.35	.05	−.09
	∞	.06	.05	−.02	.32	.06	−.04	.44	.07	−.04
		.03	.05	−.03	.16	.05	−.06	.32	.05	−.08
.5	30	.07	.38	−.07	.35	.39	−.14	.45	.42	−.16
		.03	.32	−.06	.15	.31	−.13	.30	.30	−.18
	50	.07	.37	−.06	.35	.38	−.14	.47	.41	−.15
		.02	.31	−.05	.15	.30	−.13	.31	.29	−.18
	100	.07	.36	−.07	.35	.37	−.14	.48	.41	−.15
		.02	.30	−.05	.15	.29	−.13	.31	.28	−.18
	200	.07	.35	−.07	.35	.37	−.14	.49	.41	−.15
		.02	.30	−.05	.14	.30	−.14	.30	.27	−.17
	∞	.06	.30	−.05	.32	.32	−.10	.44	.37	−.10
		.02	.29	−.05	.14	.28	−.12	.29	.27	−.17
.8	30	.07	.69	−.06	.32	.69	−.15	.40	.64	−.17
		.02	.49	−.05	.11	.48	−.12	.24	.46	−.17
	50	.06	.70	−.06	.32	.70	−.16	.42	.64	−.17
		.02	.53	−.05	.11	.54	−.13	.24	.50	−.18
	100	.06	.67	−.07	.32	.68	−.17	.45	.65	−.17
		.02	.54	−.06	.11	.55	−.13	.24	.51	−.18
	200	.06	.67	−.06	.32	.68	−.16	.46	.64	−.18
		.02	.53	−.06	.11	.55	−.12	.24	.50	−.18
	∞	.06	.63	−.06	.32	.63	−.12	.44	.62	−.13
		.02	.53	−.05	.11	.52	−.12	.23	.49	−.17

Tables 1 and 2 contain simulated biases and covariances of both moment and maximum likelihood estimators, respectively. The rows listed above correspond to the moment estimators and below to the maximum likelihood estimators. Actually, the numerical results on variances and covariances summarized in Table 2 are $n \times$ covariance, such that they are comparable. The rows labelled ∞ correspond to the asymptotic calculations, but we do not know any numerical results of the maximum likelihood estimators. Here we give the inverse of the observed information, which is superior to the expected information as an estimator of variance (see Efron and Hinkley (1978)).

The results provide evidence that both estimations are asymptotically unbiased. Even for small samples, the simulated biases are lower. The maximum bias for the moment estimation is not greater than two units in the second decimal place and for the maximum likelihood estimation is not greater than one unit in the second decimal place. The biases are all positive, but the biases of both estimations of α are negative for large α. The maximum likelihood estimator $\tilde{\alpha}$ is nearly unbiased for lower α.

It seems from Table 2 that the behaviour of both estimations is approximated by asymptotic theory for sample sizes of $n = 30$ or larger. The variances of maximum likelihood estimators are lower than moment estimators. The covariances between estimators of the dependence parameter for the two procedures are insignificantly different for all combinations of values of α, β and n.

6. Discussion and conclusions

In this paper we have considered the multivariate extreme value distribution in a nested logistic model. The aim is to illustrate the dependence structure of the model and to give a moment estimation of the dependence parameters. A suitable transformation leads to mixed independence of the variates. The transformation (2.4) is the analogue in Shi (1995*b*, 1995*c*). We believe this transformation to be very useful in multivariate extreme value analysis. It should be apparent from the simulated results that asymptotic theory of both maximum likelihood estimation and moment estimation may provide an adequate approximation for quite small samples. The relationship between parameters and moments takes a simpler form only in Gumbel marginals and, for that reason, the moment estimation is a convenient tool for the nested logistic model (1.2). We conclude that the moment method is good for all practical purposes.

Although this paper is restricted to the case of $p = 3$, the mixed independence discussed in Section 2 is a general property. For higher-dimensional models, for example, in the case of $p = 4$, based on different problems, we may have different distribution forms

$$\exp[-\{(e^{-x_1/\alpha\beta} + e^{-x_2/\alpha\beta} + e^{-x_3/\alpha\beta})^{\beta} + e^{-x_4/\alpha}\}^{\alpha}],$$
$$\exp\{-[\{(e^{-x_1/\alpha\beta\gamma} + e^{-x_2/\alpha\beta\gamma})^{\gamma} + e^{-x_3/\alpha\beta}\}^{\beta} + e^{-x_4/\alpha}]^{\alpha}\},$$
$$\exp[-\{(e^{-x_1/\alpha\beta} + e^{-x_2/\alpha\beta})^{\beta} + (e^{-x_3/\alpha\gamma} + e^{-x_4/\alpha\gamma})^{\gamma}\}^{\alpha}],$$

where $0 \leq \alpha, \beta, \gamma \leq 1$ are dependence parameters. Analogous results may be obtained in a similar way, but will have more complicated forms.

Acknowledgements

We would like to thank the Editor and referees for thorough reviews of this paper and for many helpful comments and suggestions.

REFERENCES

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions*, National Bureau of Standards, Washington D.C. (Reprinted by Dover, New York).

Coles, S. G. and Tawn, J. A. (1991). Modelling extreme multivariate events, *J. Roy. Statist. Soc. Ser. B*, **53**, 377–392.

Coles, S. G. and Tawn, J. A. (1994). Statistical method for multivariate extremes: an application to structural design, *Appl. Statist.*, **43**, 1–48.

Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: observed versus expected Fisher information, *Biometrika*, **65**, 457–487.

Kendall, M. G. and Stuart, A. (1969). *The Advanced Theory of Statistics*, Vol. 1, Charles Griffin, London.

McFadden, D. (1978). Modelling the choice of residential location, *Spatial Interaction Theory and Planning Models* (eds. A. Karlqvist, L. Lundquist, F. Snickers and J. Weibull), 75–96, North-Holland, Amsterdam.

Oakes, D. and Manatunga, A. K. (1992). Fisher information for a bivariate extreme value distribution, *Biometrika*, **79**, 827–832.

Shi, D. (1995*a*). Moment estimation for multivariate extreme value distribution, *Appl. Math. J. Chinese Univ. Ser. B*, **10**, 61–68.

Shi, D. (1995*b*). Fisher information for a multivariate extreme value distribution, *Biometrika*, **82**, 644–649.

Shi, D. (1995*c*). Multivariate extreme value distribution and its Fisher information matrix, *Acta Math. Appl. Sinica*, **11**, 421–428.

Shi, D., Smith, R. L. and Coles, S. G. (1992). Joint versus marginal estimation for bivariate extremes, Tech. Report 2074, Department of Statistics, University of North Carolina at Chapel Hill.

Shi, D., Ruan, M. and Wang, Y. (1997). Sampling method of random vector from multivariate extreme value distribution, *Chinese J. Appl. Probab. Statist.*, **13**, 75–80 (in Chinese).

Tawn, J. A. (1988). Bivariate extremes value theory: models and estimation, *Biometrika*, **75**, 397–415.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions, *Biometrika*, **77**, 245–253.

Tiago de Oliveira, J. (1980). Bivariate extremes: foundations and statistics, *Multivariate Analysis V* (ed. P. R. Krishnaiah), 349–366, North-Holland, Amsterdam.