Estimators of Bivariate Extreme Value Copulas

SUZUKAWA Akio*

1. Introduction

Copulas are functions that join multivariate distribution functions to their one-dimensional margins. A class of copulas derived from the limiting behavior of component-wise maxima of independent, identically distributed samples is that of extreme value copulas. Unlike the univariate case, there is no finite-dimensional parametrization in the multivariate extreme value distributions. In other words, the class of extreme value copulas cannot be represented by finite-dimensional parameters. The multivariate extreme value distributions have been discussed in many textbooks of extreme value theory or copulas theory, for example, Galambos [8], Resnick [20], Joe [13], Kotz and Nadarajah [15], Beirlant *et al.* [1], Castillo *et al.* [4], de Haan and Ferreira [10] and Nelsen [18].

In this paper we consider bivariate extreme value distributions. Without loss of generality, we can assume that marginal distributions are exponentials with unit means. Let X and Y be random variables with survival functions $\bar{F}(x) = P(X > x)$ and $\bar{G}(y) = P(Y > y)$, respectively. When (X, Y) follows a bivariate extreme value distribution, its joint survival function can be represented as

$$S(x, y) = \exp\left\{-(x+y)A\left(\frac{y}{x+y}\right)\right\} \tag{1}$$

for $0 \le x, y < \infty$ with $x+y > 0$, where $A : [0,1] \to [1/2, 1]$ is a convex function satisfying $A(0) = A(1) = 1$ and

$$\max(1-t, t) \le A(t) \le 1, \quad t \in [0,1]. \tag{2}$$

The representation (1) was obtained by Pickands [19], and the function A is called *Pickands dependence function.* The survival copula corresponding to the survival function S is given by

$$C(u, v) = S(-\log u, -\log v) = \exp\left[\log(uv)A\left\{\frac{\log v}{\log(uv)}\right\}\right], \quad 0 \le u, v \le 1.$$

The copula is determined by the Pickands dependence function A. Important examples of A are the lower and upper bounds of (2). If $A(t) \equiv 1$ (the upper

* Professor of Graduate School of Public Policy, Hokkaido University
suzukawa@econ.hokudai.ac.jp

bound), then X and Y are independent. If $A(t) = \max(1-t, t)$ (the lower bound), then X and Y are completely dependent, that is, $X = Y$ holds with probability one.

Several parametric models for A have been presented by Tawn [22], Joe [13], Kotz and Nadarajah [15] and Beirlant *et al.* [1]. One of the classical parametric models is the so-called logistic model, proposed by Gumbel [9], and defined by

$$A(t) = \{(1-t)^q + t^q\}^{1/q}, \quad t \in [0, 1], \tag{3}$$

where $q \geq 1$ is a dependence parameter. Independence and complete dependence correspond to $q = 1$ and $q = \infty$, respectively. Another classical parametric model is the so-called Marshall and Olkin [17] model, defined by

$$A(t) = \max(1 - \theta t, 1 - \theta(1-t)), \quad t \in [0, 1], \tag{4}$$

where $\theta \in [0, 1]$ is a dependence parameter. Independence and complete dependence correspond to $\theta = 0$ and $\theta = 1$, respectively.

We are concerned with nonparametric estimation of the Pickands dependence function A. A nonparametric estimator of A was proposed by Pickands [19]. Modifications of the Pickands estimator were suggested by Tiago de Oliveira [23], Deheuvels and Tiago de Oliveira [6], Deheuvels [5] and Hall and Tajvidi [11]. Another type of nonparametric estimator was proposed by Capéraà *et al.* [3] . Modifications to satisfy the constraints of convexity were suggested by Hall and Tajvidi [11], Jiménz, Villa-Diharce and Flores [12] and Fils-Villetard *et al.* [7].

The nonparametric estimators can be classified into two families: Pickands [19] type estimators and Capéraà-Fougères-Genest (CFG) [3] type estimators. Based on a simulation study, Capéraà *et al.* [3] discussed comparison between the Pickands-type estimators and the CFG-type estimators. The results indicate that the CFG-type estimators are preferable to the Pickands-type estimators under a wide range of dependence structures. Segers [21] gave a unified treatment of the Pickands-type and CFG-type, and showed that the CFG-type is asymptotically more efficient than Pickands-type under independence of X and Y. The unified treatment and the moment formulas obtained by Segers [21] are very useful to discuss several properties of the nonparametric estimators. In this paper, we develop the unified treatment a little further and propose a new class of nonparametric estimators for the Pickands dependence function. The class includes both the Pickands-type and the CFG-type estimators.

The outline of the paper is as follows. In Section 2, we define a new class of nonparametric estimators for the Pickands dependence function. Asymptotic properties of the nonparametric estimators are investigated in Section 3. In Section 4, we discuss asymptotic comparison of the estimators under the

Marshall-Olkin model. In Section 5, we report simulation results under the logistic model and the the Marshall-Olkin model. All the proofs are gathered in an Appendix.

2. A class of estimators for Pickands dependence function

Let (X_i, Y_i), $i = 1, 2, \ldots, n$ be random samples from the bivariate survival function (1). For $t \in [0, 1]$, define

$$\xi_i(t) = \min\left(\frac{X_i}{1-t}, \frac{Y_i}{t}\right), \; i = 1, 2, \ldots, n.$$

Then, $\xi_i(t)$ is exponentially distributed with mean

$$\frac{1}{A(t)} = E[\xi_i(t)]. \tag{5}$$

Pickands [19] estimator is an empirical version of this moment equation, and which is defined by

$$\frac{1}{\hat{A}^{\mathrm{P}}(t)} = n^{-1}\sum_{i=1}^{n} \xi_i(t) = n^{-1}\sum_{i=1}^{n} \min\left(\frac{X_i}{1-t}, \frac{Y_i}{t}\right).$$

The estimator does not satisfy the constraints $A(0) = A(1) = 1$.

Deheuvels [5] proposed an estimator $\hat{A}^{\mathrm{D}}(t)$ defined by

$$\frac{1}{\hat{A}^{\mathrm{D}}(t)} = n^{-1}\sum_{i=1}^{n}\{\xi_i(t) - (1-t)(X_i - 1) - t(Y_i - 1)\},$$

and which satisfies $\hat{A}^{\mathrm{D}}(0) = \hat{A}^{\mathrm{D}}(1) = 1$. This estimator can be considered as an empirical version of a moment equation

$$\frac{1}{A(t)} = E[\xi_i(t) - (1-t)(X_i - 1) - t(Y_i - 1)]. \tag{6}$$

Another nonparametric estimator was proposed by Capéraà *et al.* [3]. They focused on equations

$$\begin{aligned} \log A(t) &= \int_0^t \frac{\Pr\{Y_i/(X_i+Y_i) \le z\} - z}{z(1-z)} dz \\ &= -\int_t^1 \frac{\Pr\{Y_i/(X_i+Y_i) \le z\} - z}{z(1-z)} dz. \end{aligned}$$

Replacing the distribution function $\Pr\{Y_i/(X_i+Y_i) \le z\}$ by the corresponding empirical distribution function $n^{-1}\sum_{i=1}^{n} I\{Y_i/(X_i+Y_i) \le z\}$, where I is the indicator function, Capéraà, Fougères and Genest (CFG) [3] estimator is defined by

$$\log \hat{A}^{\mathrm{CFG}}(t) = p(t) \int_0^t \frac{n^{-1}\sum_{i=1}^n I\{Y_i/(X_i+Y_i) \le z\} - z}{z(1-z)} dz$$
$$-\{1-p(t)\} \int_t^1 \frac{n^{-1}\sum_{i=1}^n I\{Y_i/(X_i+Y_i) \le z\} - z}{z(1-z)} dz,$$

where $p(t)$ is an appropriate weight function on $[0,1]$. Beirlant *et al.* [1] and Segers [21] showed that $\hat{A}^{\mathrm{CFG}}(t)$ is an empirical version of an equation

$$-\log A(t) = E\left[\log \xi_i(t) - p(t)\log X_i - \{1-p(t)\}\log Y_i\right], \qquad (7)$$

and $\hat{A}^{\mathrm{CFG}}(t)$ can be expressed as

$$-\log \hat{A}^{\mathrm{CFG}}(t) = n^{-1}\sum_{i=1}^n \left[\log \xi_i(t) - p(t)\log X_i - \{1-p(t)\}\log Y_i\right].$$

In this paper, we consider Box and Cox [2] power-transformation on $[0,\infty)$ defined by

$$\varphi_\lambda(x) = \begin{cases} \lambda^{-1}(x^\lambda - 1), & \lambda > 0, \\ \log x, & \lambda = 0. \end{cases}$$

It can be easily verified that

$$E[\varphi_\lambda\{\xi_i(t)\}] = \begin{cases} \Gamma(1+\lambda)\varphi_\lambda\{1/A(t)\} + \lambda^{-1}\{\Gamma(1+\lambda)-1\}, & \lambda > 0, \\ -\log A(t) - \gamma, & \lambda = 0, \end{cases}$$

where Γ is the gamma function and γ is the Euler's constant. Let $a(t)$ and $b(t)$ be appropriate weight functions on $[0,1]$. Then, it can be seen that

$$E[\varphi_\lambda\{\xi_i(t)\} - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i)]$$
$$= \begin{cases} \Gamma(1+\lambda)\varphi_\lambda\{1/A(t)\} + \lambda^{-1}\{\Gamma(1+\lambda)-1\}\{1-a(t)-b(t)\}, & \lambda > 0, \\ -\varphi_0\{1/A(t)\} - \gamma\{1-a(t)-b(t)\}, & \lambda = 0. \end{cases}$$

From this, we can obtain an expression

$$\Gamma(1+\lambda)\varphi_\lambda\{1/A(t)\} = E[\varphi_\lambda\{\xi_i(t)\} - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i)] \qquad (8)$$
$$-\lambda^{-1}\{\Gamma(1+\lambda)-1\}\{1-a(t)-b(t)\}.$$

This equation is a generalization of equations (5), (6) and (7). When $\lambda = 1$ and $a(t) = b(t) \equiv 0$, the equation (8) reduces to (5). If $\lambda = 1$, $a(t) = 1-t$ and $b(t) = t$, then (8) gives (6). If two weight functions are chosen as $a(t) = 1-p(t)$, $b(t) = p(t)$ and let $\lambda \to 0$ in (8), then (7) is obtained.

The equation (8) suggests estimating $A(t)$ by replacing the expectation term in (8) by the corresponding sample average. We define an estimator $\hat{A}_\lambda(t; a(t), b(t))$ by

$$\begin{aligned}&\varphi_\lambda\{1/\hat{A}_\lambda(t; a(t), b(t))\} \qquad (9)\\ &= \frac{1}{n\Gamma(1+\lambda)} \sum_{i=1}^{n} [\varphi_\lambda\{\xi_i(t)\} - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i)]\\ &\quad -c_\lambda\{1 - a(t) - b(t)\},\end{aligned}$$

where

$$c_\lambda = \begin{cases} \lambda^{-1}\{\Gamma(1+\lambda) - 1\}/\Gamma(1+\lambda), & \lambda > 0, \\ -\gamma, & \lambda = 0. \end{cases}$$

Noting that $\lim_{\lambda\to 0} c_\lambda = -\gamma = c_0$, as $\lambda \to 0$ in (9), we have a CFG-type estimator

$$\begin{aligned} -\log \hat{A}_0(t; a(t), b(t)) &= \frac{1}{n} \sum_{i=1}^{n} [\log \xi_i(t) - a(t) \log X_i - b(t) \log Y_i] \\ &\quad +\gamma\{1 - a(t) - b(t)\}. \end{aligned}$$

From (8) and (9), we have

$$\begin{aligned}&E\left[\varphi_\lambda\{1/\hat{A}_\lambda(t; a(t), b(t))\}\right]\\ &= \frac{1}{\Gamma(1+\lambda)} E\left[\varphi_\lambda\{\xi_i(t)\} - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i)\right] - c_\lambda\{1 - a(t) - b(t)\}\\ &= \varphi_\lambda\{1/A(t)\},\end{aligned}$$

for any fixed $t \in [0, 1]$ and $\lambda \geq 0$. Hence, $\varphi_\lambda\{1/\hat{A}_\lambda(t; a(t), b(t))\}$ is an unbiased estimator of $\varphi_\lambda\{1/A(t)\}$. In (9), the term $-c_\lambda\{1 - a(t) - b(t)\}$ is a bias-correction term. If $a(t) + b(t) = 1$, this term is not needed.

The Pickands [19] estimator, the Deheuvels [5] estimator and the Capéraà, Fougères and Genest (CFG) [3] estimator can be expressed as

$$\hat{A}^{\mathrm{P}}(t) = \hat{A}_1(t; 0, 0), \quad \hat{A}^{\mathrm{D}}(t) = \hat{A}_1(t; 1-t, t), \quad \hat{A}^{\mathrm{CFG}}(t) = \hat{A}_0(t; 1-p(t), p(t)).$$

In the next section, we investigate asymptotic properties of $\hat{A}_\lambda(t; a(t), b(t))$ with $\lambda \geq 0$.

3. Asymptotic Properties

Assume that weight functions a and b are bounded on $[0, 1]$. Then, for any fixed $t \in [0, 1]$ and $\lambda \geq 0$,

$$E[|\varphi_\lambda\{\xi_i(t)\} - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i)|] < \infty.$$

Thus, by the strong law of large numbers, for any fixed $t \in [0,1]$ and $\lambda \geq 0$,

$$\lim_{n\to\infty} \varphi_\lambda\{1/\hat{A}_\lambda(t; a(t), b(t))\} = \varphi_\lambda\{1/A(t)\} \quad \text{a.s..}$$

Since the transformation φ_λ is continuous, $\hat{A}_\lambda(t; a(t), b(t))$ is consistent, that is,

$$\lim_{n\to\infty} \hat{A}_\lambda(t; a(t), b(t)) = A(t) \quad \text{a.s..}$$

The second moment of the transformed variate $\varphi_\lambda\{\xi_i(\cdot)\}$ is given by the following lemma.

Lemma 1. *For $\lambda > 0$ and $0 \leq s \leq t \leq 1$, covariance between $\varphi_\lambda\{\xi(s)\}$ and $\varphi_\lambda\{\xi(t)\}$ is given by*

$$\begin{aligned}
&\mathrm{Cov}\left[\varphi_\lambda\{\xi(s)\}, \varphi_\lambda\{\xi(t)\}\right] \\
&= \frac{\Gamma(1+2\lambda)}{2\lambda^2}\left[\left[\frac{1-t}{(1-s)\{A(t)\}^2}\right]^\lambda + \left[\frac{s}{t\{A(s)\}^2}\right]^\lambda \right. \\
&\quad \left. + \frac{\lambda}{(1-s)^\lambda t^\lambda}\int_s^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}dw\right] - \frac{\{\Gamma(1+\lambda)\}^2}{\lambda^2\{A(s)A(t)\}^\lambda}.
\end{aligned}$$

Letting $\lambda = 1$ in Lemma 1, we have

$$\begin{aligned}
&\mathrm{Cov}\left[\varphi_1\{\xi(s)\}, \varphi_1\{\xi(t)\}\right] = \mathrm{Cov}\left[\xi(s), \xi(t)\right] \\
&= \frac{1-t}{(1-s)\{A(t)\}^2} + \frac{s}{t\{A(s)\}^2} + \frac{1}{(1-s)t}\int_s^t \frac{1}{\{A(w)\}^2}dw - \frac{1}{A(s)A(t)}.
\end{aligned}$$

This formula has been obtained in Theorem 1 of Segers [21]. Covariance formula between $\varphi_0\{\xi(s)\} = \log\xi(s)$ and $\varphi_0\{\xi(t)\} = \log\xi(t)$ has been given in Theorem 2 of Segers [21]. The formula can be also derived from Lemma 1 as $\lambda \to 0$. The result is given in the next corollary.

Corollary 2. *For $0 \leq s \leq t \leq 1$, covariance between $\varphi_0\{\xi(s)\} = \log\xi(s)$ and $\varphi_0\{\xi(t)\} = \log\xi(t)$ is given by*

$$\begin{aligned}
&\mathrm{Cov}[\varphi_0\{\xi(s)\}, \varphi_0\{\xi(t)\}] = \lim_{\lambda\to 0}\mathrm{Cov}[\varphi_\lambda\{\xi(s)\}, \varphi_\lambda\{\xi(t)\}] \\
&= \frac{\pi^2}{6} + (\log t)\log\frac{1-t}{1-s} + \int_s^t \frac{\log w}{1-w}dw - \{\log A(t)\}\log\frac{1-t}{1-s} \\
&\quad -\{\log A(s)\}\log\frac{s}{t} + \frac{1}{2}\left\{\log\frac{A(s)}{A(t)}\right\}^2 - \int_s^t \frac{\log A(w)}{w(1-w)}dw.
\end{aligned}$$

The estimator $\hat{A}_\lambda(t; a(t), b(t))$ consists of random samples of 3-dimensional random vector

$$[\varphi_\lambda\{\xi(t)\}, \varphi_\lambda(X), \varphi_\lambda(Y)].$$

A covariance matrix of it is directly obtained from Lemma 1.

Corollary 3. *For any fixed $t \in [0,1]$ and $\lambda > 0$, a covariance matrix of 3-dimensional random vector $[\varphi_\lambda\{\xi(t)\}, \varphi_\lambda(X), \varphi_\lambda(Y)]$ is given by*

$$\mathbf{\Sigma}_\lambda(t) = \begin{bmatrix} \sigma_\lambda(t,t) & \sigma_\lambda(0,t) & \sigma_\lambda(t,1) \\ \sigma_\lambda(0,t) & \sigma_\lambda(0,0) & \sigma_\lambda(0,1) \\ \sigma_\lambda(t,1) & \sigma_\lambda(0,1) & \sigma_\lambda(1,1) \end{bmatrix}, \tag{10}$$

where

$$\begin{aligned}
\sigma_\lambda(t,t) &= \mathrm{Var}[\varphi_\lambda\{\xi(t)\}] = \frac{\Gamma(1+2\lambda) - \{\Gamma(1+\lambda)\}^2}{\lambda^2\{A(t)\}^{2\lambda}}, \\
\sigma_\lambda(0,0) &= \sigma_\lambda(1,1) = \mathrm{Var}[\varphi_\lambda(X)] = \mathrm{Var}[\varphi_\lambda(Y)] \\
&= \left[\Gamma(1+2\lambda) - \{\Gamma(1+\lambda)\}^2\right]/\lambda^2, \\
\sigma_\lambda(0,t) &= \mathrm{Cov}[\varphi_\lambda(X), \varphi_\lambda\{\xi(t)\}] \\
&= \frac{\Gamma(1+2\lambda)}{2\lambda^2}\left[\left[\frac{1-t}{\{A(t)\}^2}\right]^\lambda + \frac{\lambda}{t^\lambda}\int_0^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}dw\right] \\
&\quad - \frac{\{\Gamma(1+\lambda)\}^2}{\lambda^2\{A(t)\}^\lambda}, \\
\sigma_\lambda(t,1) &= \mathrm{Cov}[\varphi_\lambda\{\xi(t)\}, \varphi_\lambda(Y)] \\
&= \frac{\Gamma(1+2\lambda)}{2\lambda^2}\left[\left[\frac{t}{\{A(t)\}^2}\right]^\lambda + \frac{\lambda}{(1-t)^\lambda}\int_t^1 \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}dw\right] \\
&\quad - \frac{\{\Gamma(1+\lambda)\}^2}{\lambda^2\{A(t)\}^\lambda}, \\
\sigma_\lambda(0,1) &= \mathrm{Cov}[\varphi_\lambda(X), \varphi_\lambda(Y)] \\
&= \frac{\Gamma(1+2\lambda)}{2\lambda}\int_0^1 \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}dw - \frac{\{\Gamma(1+\lambda)\}^2}{\lambda^2}.
\end{aligned}$$

Tawn [22] has stated that correlation between X and Y, whose margins are exponential, is given by

$$\mathrm{Cor}[X,Y] = \int_0^1 \frac{1}{\{A(w)\}^2}dw - 1.$$

From Corollary 3, we can obtain a generalized formula

$$\begin{aligned}
\mathrm{Cor}[X^\lambda, Y^\lambda] &= \mathrm{Cor}[\varphi_\lambda(X), \varphi_\lambda(Y)] = \sigma_\lambda(0,1)/\sigma_\lambda(0,0) \\
&= \frac{2^{-1}\lambda\Gamma(1+2\lambda)\int_0^1 \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}dw - \{\Gamma(1+\lambda)\}^2}{\Gamma(1+2\lambda) - \{\Gamma(1+\lambda)\}^2}.
\end{aligned}$$

This is a correlation under Weibull margins. For any $\lambda > 0$, the two extreme case $\mathrm{Cor}[X^\lambda, Y^\lambda] = 0$ and 1 correspond to independence $A(w) \equiv 1$ and complete dependence $A(w) = \max(1-w, w)$.

Fundamental asymptotic properties of $\hat{A}_\lambda(t; a(t), b(t))$ are given by the following theorem.

Theorem 4. *Assume that weight functions a and b are bounded on $[0,1]$. Then, for any fixed $t \in [0,1]$ and $\lambda \geq 0$,*

$$E\left[\hat{A}_\lambda(t; a(t), b(t))\right] = A(t) + \frac{(1+\lambda)\{A(t)\}^{2\lambda+1}\ \tau_\lambda^2(t; a(t), b(t))}{2n} + O\left(\frac{1}{n^2}\right),$$

$$\mathrm{Var}\left[\hat{A}_\lambda(t; a(t), b(t))\right] = \frac{\{A(t)\}^{2(\lambda+1)}\ \tau_\lambda^2(t; a(t), b(t))}{n} + O\left(\frac{1}{n^2}\right)$$

and

$$\sqrt{n}\left\{\hat{A}_\lambda(t; a(t), b(t)) - A(t)\right\} \xrightarrow{L} N\left(0, \{A(t)\}^{2(\lambda+1)}\ \tau_\lambda^2(t; a(t), b(t))\right),$$

where

$$\tau_\lambda^2(t; a(t), b(t)) = \frac{\boldsymbol{\eta}'(t)\boldsymbol{\Sigma}_\lambda(t)\boldsymbol{\eta}(t)}{\{\Gamma(1+\lambda)\}^2},$$

$\boldsymbol{\eta}(t)$ is a 3-dimensional column vector defined by $\boldsymbol{\eta}'(t) = [1, -a(t), -b(t)]$ and $\boldsymbol{\Sigma}_\lambda(t)$ is the covariance matrix defined in Corollary 3.

From Theorem 4, the expected squared error of $\hat{A}_\lambda(t; a(t), b(t))$ is given by

$$E\left[\left\{\hat{A}_\lambda(t; a(t), b(t)) - A(t)\right\}^2\right] = \mathrm{Var}\left[\hat{A}_\lambda(t; a(t), b(t))\right] + O\left(\frac{1}{n^2}\right),$$

and its main contribution comes from the variance term.

When $a(t) = b(t) \equiv 0$, the asymptotic variance is

$$\{A(t)\}^{2(\lambda+1)}\ \tau_\lambda^2(t; 0, 0) = \frac{\sigma_\lambda(0,0)}{\{\Gamma(1+\lambda)\}^2}\{A(t)\}^2.$$

Thus, asymptotic relative efficiency (ARE) of $\hat{A}_\lambda(t; 0, 0)$ with respect to the Pickands estimator $\hat{A}^{\mathrm{P}}(t) = \hat{A}_1(t; 0, 0)$ is given by

$$\mathrm{ARE}\left[\hat{A}_\lambda(t; 0, 0), \hat{A}^{\mathrm{P}}(t)\right] = \frac{\{\Gamma(1+\lambda)\}^2}{\sigma_\lambda(0,0)},$$

which does not depend on t. Figure 1 shows the ARE for $\lambda \geq 0$. We can see that the ARE is less than one for $\lambda \neq 1$. If weight functions are not used, then the Pickands estimator is asymptotically preferable. The estimator $\hat{A}_0(t; 0, 0)$ is an CFG-type estimator without weight functions, and its ARE is $6/\pi^2 \approx 0.608$. From these results, it can be seen that the weight functions $a(t)$ and $b(t)$ play a very important role in $\hat{A}_\lambda(t; a(t), b(t))$.

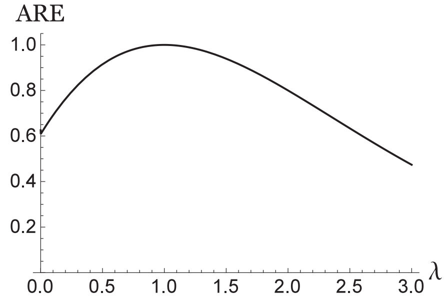

Figure 1. Asymptotic relative efficiency of $\hat{A}_\lambda(t;0,0)$ with respect to $\hat{A}^{\mathrm{P}}(t) = \hat{A}_1(t;0,0)$.

We now consider some conditions for the weight functions. From Theorem 4, asymptotic variances of $\hat{A}_\lambda(0;a(0),b(0))$ and $\hat{A}_\lambda(1;a(1),b(1))$ are given by

$$\tau^2_\lambda(0;a(0),b(0)) = \frac{\{(1-a(0))^2+(b(0))^2\}\sigma_\lambda(0,0)-2b(0)(1-a(0))\sigma_\lambda(0,1)}{\{\Gamma(1+\lambda)\}^2},$$
$$\tau^2_\lambda(1;a(1),b(1)) = \frac{\{(a(1))^2+(1-b(1))^2\}\sigma_\lambda(1,1)-2a(1)(1-b(1))\sigma_\lambda(0,1)}{\{\Gamma(1+\lambda)\}^2},$$

respectively. Thus, if $a(t)$ and $b(t)$ satisfy conditions $a(0) = b(1) = 1$ and $a(1) = b(0) = 0$, then the above two variances vanish. This is natural because, under these conditions, $\hat{A}_\lambda(t;a(t),b(t))$ satisfies a preferable property

$$\hat{A}_\lambda(0;a(0),b(0)) = \hat{A}_\lambda(1;a(1),b(1)) = 1.$$

Partition $\boldsymbol{\Sigma}_\lambda(t)$ of (10) as

$$\boldsymbol{\Sigma}_\lambda(t) = \left[\begin{array}{c|cc} \sigma_\lambda(t,t) & \sigma_\lambda(0,t) & \sigma_\lambda(t,1) \\ \hline \sigma_\lambda(0,t) & \sigma_\lambda(0,0) & \sigma_\lambda(0,1) \\ \sigma_\lambda(t,1) & \sigma_\lambda(0,1) & \sigma_\lambda(1,1) \end{array}\right] = \left[\begin{array}{cc} \sigma_\lambda(t,t) & \tilde{\boldsymbol{\sigma}}'_\lambda(t) \\ \tilde{\boldsymbol{\sigma}}_\lambda(t) & \tilde{\boldsymbol{\Sigma}}_\lambda \end{array}\right]$$

and put $\tilde{\boldsymbol{\eta}}'(t) = [a(t), b(t)]$ Then, the asymptotic variance of $\hat{A}_\lambda(t;a(t),b(t))$ can be expressed as

$$\frac{\{A(t)\}^{2(1+\lambda)}}{\{\Gamma(1+\lambda)\}^2}\left\{\sigma_\lambda(t,t) - 2\tilde{\boldsymbol{\sigma}}'_\lambda(t)\tilde{\boldsymbol{\eta}}(t) + \tilde{\boldsymbol{\eta}}'(t)\tilde{\boldsymbol{\Sigma}}_\lambda\tilde{\boldsymbol{\eta}}(t)\right\}.$$

For any fixed $t \in [0,1]$ and $\lambda \geq 0$, this is minimized at

$$\tilde{\boldsymbol{\eta}}(t) = \left[\begin{array}{c} a^*_\lambda(t) \\ b^*_\lambda(t) \end{array}\right] = \tilde{\boldsymbol{\Sigma}}^{-1}_\lambda \tilde{\boldsymbol{\sigma}}_\lambda(t) = \left[\begin{array}{cc} \sigma_\lambda(0,0) & \sigma_\lambda(0,1) \\ \sigma_\lambda(0,1) & \sigma_\lambda(1,1) \end{array}\right]^{-1} \left[\begin{array}{c} \sigma_\lambda(0,t) \\ \sigma_\lambda(t,1) \end{array}\right] \qquad (11)$$

if $\widetilde{\boldsymbol{\Sigma}}_\lambda$ is nonsingular. Corresponding minimal variance is

$$\frac{\{A(t)\}^{2(1+\lambda)}}{\{\Gamma(1+\lambda)\}^2}\left\{\sigma_\lambda(t,t)-\tilde{\boldsymbol{\sigma}}'_\lambda(t)\widetilde{\boldsymbol{\Sigma}}_\lambda^{-1}\tilde{\boldsymbol{\sigma}}(t)\right\}. \tag{12}$$

We call $a^*_\lambda(t)$ and $b^*_\lambda(t)$ defined by (11) as optimal weight functions. These functions satisfy $a^*_\lambda(0) = b^*_\lambda(1) = 1$ and $a^*_\lambda(1) = b^*_\lambda(0) = 0$, and hence, $\hat{A}_\lambda(t; a^*_\lambda(t), b^*_\lambda(t))$ satisfy

$$\hat{A}_\lambda(0; a^*(0), b^*(0)) = \hat{A}_\lambda(1; a^*(1), b^*(1)) = 1.$$

Singularity of $\widetilde{\boldsymbol{\Sigma}}_\lambda$ occurs only in complete dependence case because $\sigma_\lambda(0,1) = \sigma_\lambda(0,0)$ if and only if $A(w) = \max(1-w, w)$ for all $w \in [0,1]$. In other words, except for complete dependence case, the optimal weight functions are uniquely determined by (11). However, we can not know the optimal weight functions because they depend on the unknown dependence function A.

4. Asymptotic Comparison under the Marshall-Olkin model

The purpose here is to explore a little further into optimalities of λ, $a(t)$ and $b(t)$ in the estimator $\hat{A}_\lambda(t; a(t), b(t))$. It seems reasonable to consider that the optimalities depend on the unknown $A(t)$. It is quite likely that strength of dependence between X and Y mainly influences the optimalities of λ, $a(t)$ and $b(t)$. In order to investigate such an influence, we assume the Marshall-Olkin model defined by (4). It is a symmetric case of the nondifferentiable asymmetric logistic model introduced by Tawn [22], and is just the Marshall and Olkin [17] bivariate exponential model transformed to have unit exponential margins. When $\theta = 0$, we have independence. Complete dependence corresponds to $\theta = 1$. The joint distribution of X and Y is singular on the line $x = y$, and

$$\mathrm{Cor}[X, Y] = \Pr[X = Y] = \theta/(2-\theta).$$

Define a function f_λ on $[0,1]$ by

$$f_\lambda(t) = \lambda t^{-\lambda}\int_0^t w^{\lambda-1}(1-w)^{\lambda-1}\,dw.$$

This is one of the Gaussian hypergeometric function, that is,

$$f_\lambda(t) = {}_2F_1[\lambda, 1-\lambda; \lambda+1; t] = \sum_{j=0}^{\infty}\frac{(\lambda)_j(1-\lambda)_j\, t^j}{(1+\lambda)_j\, j!},$$

where $(c)_j$ is Pochhammer's symbol defined by $(c)_j = c(c+1)\cdots(c+j-1)$, cf. Johnson *et al.* [14].

Under (4), it holds that

$$\lambda \int_s^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}\, dw \tag{13}$$

$$= \begin{cases} \left(\frac{t}{A(t)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)t}{A(t)}\right) - \left(\frac{s}{A(s)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)s}{A(s)}\right), & t < \frac{1}{2}, \\ \frac{2}{(2-\theta)^\lambda} f_\lambda\left(\frac{1-\theta}{2-\theta}\right) - \left(\frac{s}{A(s)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)s}{A(s)}\right) & \\ -\left(\frac{1-t}{A(t)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)(1-t)}{A(t)}\right), & s < \frac{1}{2} \le t, \\ \left(\frac{1-s}{A(s)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)(1-s)}{A(s)}\right) - \left(\frac{1-t}{A(t)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)(1-t)}{A(t)}\right), & \frac{1}{2} \le s, \end{cases}$$

for $0 \le s \le t \le 1$. A proof of this is in Appendix.

From (13) and Corollary 3, we have, under (4),

$$\frac{2\lambda^2\sigma_\lambda(0,0)}{\Gamma(1+2\lambda)} = \frac{2\lambda^2\sigma_\lambda(1,1)}{\Gamma(1+2\lambda)} = 2 - f_\lambda(1), \quad \frac{2\lambda^2\sigma_\lambda(t,t)}{\Gamma(1+2\lambda)} = \frac{2-f_\lambda(1)}{\{A(t)\}^{2\lambda}}, \tag{14}$$

$$\begin{aligned} \frac{2\lambda^2\sigma_\lambda(0,t)}{\Gamma(1+2\lambda)} &= \frac{(1-t)^\lambda}{\{A(t)\}^{2\lambda}} - \frac{f_\lambda(1)}{\{A(t)\}^\lambda} + \frac{I(t<1/2)}{\{A(t)\}^\lambda} f_\lambda\left(\frac{(1-\theta)t}{A(t)}\right) \\ &\quad + I(t \ge 1/2)\left\{\frac{2}{t^\lambda(2-\theta)^\lambda} f_\lambda\left(\frac{1-\theta}{2-\theta}\right)\right. \\ &\quad \left. - \left(\frac{1-t}{tA(t)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)(1-t)}{A(t)}\right)\right\}, \\ \frac{2\lambda^2\sigma_\lambda(t,1)}{\Gamma(1+2\lambda)} &= \frac{t^\lambda}{\{A(t)\}^{2\lambda}} - \frac{f_\lambda(1)}{\{A(t)\}^\lambda} + \frac{I(t>1/2)}{\{A(t)\}^\lambda} f_\lambda\left(\frac{(1-\theta)(1-t)}{A(t)}\right) \\ &\quad + I(t \le 1/2)\left\{\frac{2}{(1-t)^\lambda(2-\theta)^\lambda} f_\lambda\left(\frac{1-\theta}{2-\theta}\right)\right. \\ &\quad \left. - \left(\frac{t}{(1-t)A(t)}\right)^\lambda f_\lambda\left(\frac{(1-\theta)t}{A(t)}\right)\right\}, \\ \frac{2\lambda^2\sigma_\lambda(0,1)}{\Gamma(1+2\lambda)} &= \frac{2}{(2-\theta)^\lambda} f_\lambda\left(\frac{1-\theta}{2-\theta}\right) - f_\lambda(1), \end{aligned}$$

where I is the indicator function.

The correlation coefficient between X^λ and Y^λ is given by

$$\mathrm{Cor}[X^\lambda, Y^\lambda] = \mathrm{Cor}[\varphi_\lambda(X), \varphi_\lambda(Y)] = \frac{\sigma_\lambda(0,1)}{\sigma_\lambda(0,0)} = \frac{2f_\lambda\left(\frac{1-\theta}{2-\theta}\right) - (2-\theta)^\lambda f_\lambda(1)}{(2-\theta)^\lambda\{2-f_\lambda(1)\}},$$

and which is non-increasing with $\lambda \ge 0$.

Under complete dependence: $\theta = 1$

Substitute $\theta = 1$ into (14). Then, we have covariance matrix of (10) as

$$\mathbf{\Sigma}_\lambda(t) = \frac{\Gamma(1+2\lambda)\{2 - f_\lambda(1)\}}{2\lambda^2} \begin{bmatrix} \{A(t)\}^{-2\lambda} & \{A(t)\}^{-\lambda} & \{A(t)\}^{-\lambda} \\ \{A(t)\}^{-\lambda} & 1 & 1 \\ \{A(t)\}^{-\lambda} & 1 & 1 \end{bmatrix}$$

under complete dependence. Because of the singularity, the optimal weight functions are not determined by (11). In this case, the asymptotic variance of $\hat{A}_\lambda(t; a(t), b(t))$ is given by

$$\frac{\{2 - f_\lambda(1)\}\{A(t)\}^2}{\lambda^2 f_\lambda(1)} \left[1 - \{A(t)\}^\lambda \{a(t) + b(t)\}\right]^2. \tag{15}$$

Under complete dependence, if two weight functions satisfy a condition

$$a(t) + b(t) = \frac{1}{\{\max(1-t, t)\}^\lambda} \quad \text{for all } t \in [0, 1], \tag{16}$$

then, (15) vanishes. This is a natural result. If (16) holds, then we have

$$\hat{A}_\lambda(t; a(t), b(t)) = A(t) = \max(1-t, t) \quad \text{for all } t \in [0, 1] \tag{17}$$

with probability one. This is shown in Appendix. When $\lambda = 0$, (16) reduces to a simple condition $a(t) + b(t) \equiv 1$. This shows that, in CFG-type estimator ($\lambda = 0$), the weight functions $a(t) = p(t)$ and $b(t) = 1 - p(t)$ are optimal under complete dependence.

Under independence: $\theta = 0$

When $\theta = 0$, the covariance matrix of (10) is given by

$$\mathbf{\Sigma}_\lambda(t) = \frac{\Gamma(1+2\lambda)}{2\lambda^2} \times \begin{bmatrix} 2 - f_\lambda(1) & (1-t)^\lambda - f_\lambda(1) + f_\lambda(t) & t^\lambda - f_\lambda(1) + f_\lambda(1-t) \\ (1-t)^\lambda - f_\lambda(1) + f_\lambda(t) & 2 - f_\lambda(1) & 0 \\ t^\lambda - f_\lambda(1) + f_\lambda(1-t) & 0 & 2 - f_\lambda(1) \end{bmatrix}.$$

From (11) and (12), the optimal weight functions and the corresponding minimal variance are given by

$$a_\lambda^*(t) = \frac{(1-t)^\lambda - f_\lambda(1) + f_\lambda(t)}{2 - f_\lambda(1)}, \quad b_\lambda^*(t) = \frac{t^\lambda - f_\lambda(1) + f_\lambda(1-t)}{2 - f_\lambda(1)}$$

and

$$\frac{1}{\lambda^2 f_\lambda(1)} \left[2 - f_\lambda(1) - \frac{\{(1-t)^\lambda - f_\lambda(1) + f_\lambda(t)\}^2 + \{t^\lambda - f_\lambda(1) + f_\lambda(1-t)\}^2}{2 - f_\lambda(1)}\right],$$

respectively. When $\lambda = 1$, these reduce to $a_1^*(t) = 1-t$, $b_1^*(t) = t$ and $2t(1-t)$, respectively, and which have been already derived by Segers [21]. When $\lambda = 0$, the optimal weight functions and the corresponding minimal variance are given by

$$a_0^*(t) = 1 - \frac{6}{\pi^2} L_2(t), \quad b_0^*(t) = 1 - \frac{6}{\pi^2} L_2(1-t)$$

and

$$\frac{\pi^2}{6} - \frac{\{\pi^2/6 - L_2(t)\}^2 + \{\pi^2/6 - L_2(1-t)\}^2}{\pi^2/6},$$

where $L_2(t)$ is the dilogarithm function defined by

$$L_2(t) = -\int_0^t \frac{\log(1-w)}{w}\, dw = \sum_{j=1}^{\infty} \frac{t^j}{j^2}.$$

These results for $\lambda = 0$ have been also obtained by Segers [21].

For $\lambda = 0, 1/5, 1/3$ and 1, the dotted curves in Figure 2 show the optimal weight functions a_λ^* under independence. The reason that two values of λ are chosen as $1/5$ and $1/3$ will be discussed later. The optimal b_λ^* is given by $b_\lambda^*(t) = a_\lambda^*(1-t)$. The optimal weight functions under independence do not satisfy the optimality condition (16) under complete dependence. They are not optimal under complete dependence.

Under the Marshall-Olkin model: $0 \le \theta \le 1$

Covariances for general $\theta \in [0, 1]$ are given by (14). Unless $\theta = 1$ (complete dependence), the optimal weight functions $a_\lambda^*(t)$ and $b_\lambda^*(t)$ are uniquely determined by (11) and they are symmetric about $t = 1/2$. For $\lambda = 0, 1/5, 1/3$ and 1, the optimal weight functions $a_\lambda^*(t)$ are shown in Figure 2 when $\theta = 0, 0.5$ and 0.8. In the Marshall-Olkin model (4), the correlation coefficient between X and Y is given by $\theta/(2-\theta)$. The two values $\theta = 0.5$ and $\theta = 0.8$ correspond to the correlation coefficient$= 1/3$ and $2/3$, respectively. The parameter θ can be considered as one of the global dependence measures between X and Y. From Figure 2, we can see that $a_{1/5}^*(t)$ and $a_{1/3}^*(t)$ do not so influenced by θ. They are near to the simple weight function $1-t$. This is an important property. For $\lambda = 1/5$ and $1/3$, we can use the simple choice $a(t) = 1-t$ and $b(t) = t$ independently of the dependency.

For each $\lambda \ge 0$, the minimal asymptotic variance is given by (12). The minimal asymptotic variances for $\lambda = 0, 1/5, 1/3$ and 1 are shown in Figure 3. If two weight functions are optimally chosen, the Pickands-type ($\lambda = 1$) is inefficient for all θ. Under independence ($\theta = 0$), the minimal asymptotic variance of the CFG-type ($\lambda = 0$) is the smallest. On the other hand, under $\theta = 0.5$ and 0.8, the estimators of $\lambda = 1/5$ and $1/3$ have smaller minimal variances than

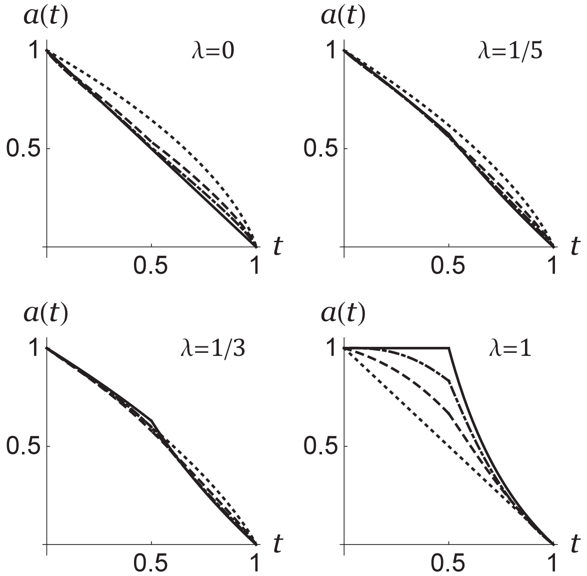

Figure 2. Optimal weight functions under Marshall and Olkin's Model (4). The dotted, broken and broken-dotted curves are under $\theta = 0$, 0.5 and 0.8, respectively. In each figure, the solid curve is the weight function $\tilde{a}_\lambda(t)$ defined by (18).

the CFG-type ($\lambda = 0$). When $\theta \geq 1/2$, the correlation coefficient between X and Y is not less than 1/3. If the weight functions are optimally chosen, the CFG-type is preferable under weak dependence asymptotically. The estimators $\lambda = 1/5$ and 1/3 are superior to the CFG-type under dependence.

When the weight functions are optimally chosen, we can obtain the asymptotic relative efficiency (ARE) of $\hat{A}_\lambda(t, a^*_\lambda(t), b^*_\lambda(t))$ with respect to the CFG-type $\hat{A}_0(t, a^*_0(t), b^*_0(t))$. Let us consider a relation between the ARE and the parameter λ. We focus on the ARE at $t = 1/4$ and 1/2. Figure 4 shows the relationship under $\theta = 0, 0.5$ and 0.8. Under independence ($\theta = 0$, the dotted curves), the ARE for any $\lambda > 0$ is less than one. On the other hand, under dependence ($\theta = 0.5$ and 0.8, the dashed and solid curves), the ARE's are greater than one if $0 < \lambda < 0.4$. Except for the independent case, there exists a optimal $\lambda \in (0, 0.4)$. This is the reason that we select $\lambda = 1/5$ and 1/3.

The above discussion is based on the optimal choice of the weight functions. However, the optimals depend on the unknown dependence function $A(t)$. Segers [21] proposed an adaptive estimator, in which an initial estimator of $A(t)$ was used for estimating the optimal weight functions. Here we do not

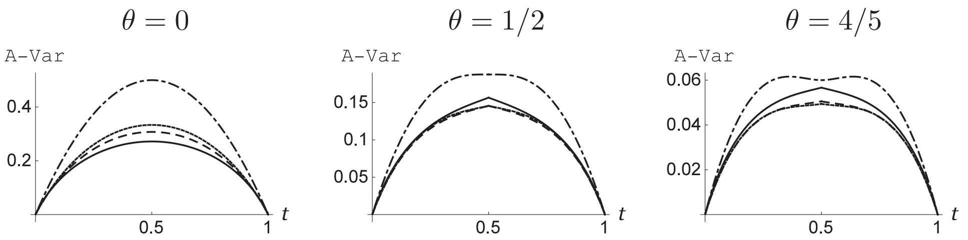

Figure 3. Asymptotic variance of $\hat{A}_\lambda(t; a^*_\lambda(t), b^*_\lambda(t))$. In each figure, the solid curve is for $\lambda = 0$, the dotted curve is for $\lambda = 1/3$, the broken curve is for $\lambda = 1/5$ and the broken-dotted curve is for $\lambda = 1$.

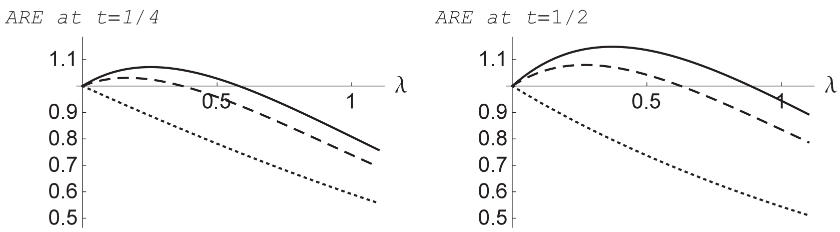

Figure 4. Asymptotic relative efficiency of $\hat{A}_\lambda(t; a^*_\lambda(t), b^*_\lambda(t))$ with respect to $\hat{A}_0(t; a^*_0(t), b^*_0(t))$. In each figure, the dotted curve, the broken curve and the solid curve are under $\theta = 0$, 0.5 and 0.8, respectively.

use such an adaptive method. We propose weight functions

$$\tilde{a}_\lambda(t) = \frac{1-t}{\{\max(1-t,t)\}^\lambda} \quad \text{and} \quad \tilde{b}_\lambda(t) = \frac{t}{\{\max(1-t,t)\}^\lambda}. \tag{18}$$

When $\lambda = 0$ (CFG-type), these are the simple weight functions $1-t$ and t, respectively. When $\lambda = 0$, the simple weights are optimal under complete dependence. For any $\lambda \geq 0$, $\tilde{a}_\lambda(t)$ and $\tilde{b}_\lambda(t)$ of (18) satisfy (16). Thus, these are optimal under complete dependence. The condition (16) is important in the sense that (17) holds if the bivariate data are completely dependent. The weight functions defined by (18) can be considered as a generalization of the simple weight functions. The solid curves in Figure 2 show $\tilde{a}_\lambda(t)$ for $\lambda = 0, 1/5, 1/3$ and 1.

Figure 5 shows the asymptotic variance of

$$\hat{A}_0(t; 1-t, t), \quad \hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t)), \quad \hat{A}_{1/3}(t; \tilde{a}_{1/3}(t), \tilde{b}_{1/3}(t)). \tag{19}$$

We can see that the estimator $\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$ is asymtotically more efficient than others. Even if $\theta = 0$, $\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$ is superior to the CFG type with the simple weights. As we have seen in Figure 3, if the weight functions can be chosen optimally for each $\lambda \geq 0$, the CFG-type ($\lambda = 0$) is more efficient under independence. As has been pointed out, the CFG-type is sensitive to choice of the weight functions. In the CFG-type, the simple weights $a(t) = 1-t$ and $b(t) = t$ are optimal under complete dependence. However, under weak dependence, the simple choice is inadequate for the CFG-type. This is a reason that $\hat{A}_0(t; 1-t, t)$ is inferior to $\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$ even if $\theta = 0$. Figure 6 shows the relationship between the parameter $\lambda(\geq 0)$ and the ARE of $\hat{A}_\lambda(t, \tilde{a}_\lambda(t), \tilde{b}_\lambda(t))$ with respect to $\hat{A}_0(t, 1-t, t)$. It can be seen that each curve representing the ARE attains a maximum around at $\lambda = 1/5$.

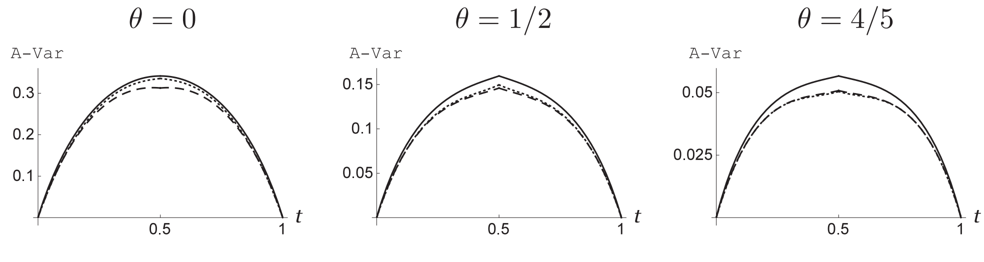

Figure 5. Asymptotic variance of $\hat{A}_\lambda(t; \tilde{a}_\lambda(t), \tilde{b}_\lambda(t))$. In each figure, the solid curve is for $\lambda = 0$, the dotted curve is for $\lambda = 1/3$ and the broken curve is for $\lambda = 1/5$.

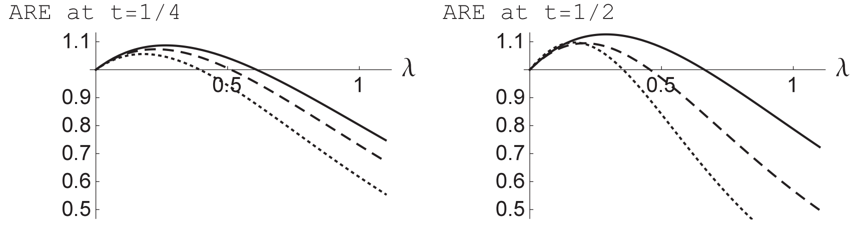

Figure 6. Asymptotic relative efficiency of $\hat{A}_\lambda(t; \tilde{a}^*_\lambda(t), \tilde{b}^*_\lambda(t))$ with respect to $\hat{A}_0(t; 1-t, t)$. In each figure, the dotted curve, the broken curve and the solid curve are under $\theta = 0$, 0.5 and 0.8, respectively.

5. Simulations

To compare the finite-sample behaviours of the estimators (19), we carried out the Monte-Carlo experoment under the logistic model (3) and the Marshall-Olkin model (4). In the logistic model, Kendall's tau and the correlation coefficient are given by

$$1 - 1/q \quad \text{and} \quad q^{-1}\{\Gamma(2/q)\}^{-1}\{\Gamma(1/q)\}^2 - 1,$$

respectively.

For sample saizes $n = 30$, and 100, data are generated. In the Marshall-Olkin model, the parameter values are selected as $\theta = 0$, $1/2$ and $4/5$. The two values $\theta = 1/2$ and $4/5$ correspond to the correlation coefficients= $1/3$ and $2/3$, respectively. In the logistic model, we choose $q = 1.435$ and $q = 2.378$, which correspond to the correlation coefficients= $1/3$ and $2/3$, respectively. In each experiment, the pointwise mean square errors (MSE) of three estimators (19) are numerically evaluated.

Under the Marshall-Olkin model, simulation results show similar features to the theoretical variance curves given by Figure 5.

Figure 7 shows the MSE under the logistic model. Under this model, $\hat{A}_{1/3}(t; \tilde{a}_{1/3}(t), \tilde{b}_{1/3}(t))$ is not preferable. The difference between $\hat{A}_0(t; 1-t, t)$ and $\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$ are slight. When $q = 1.435$, $\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$ has smaller MSE than $\hat{A}_0(t; 1-t, t)$. However, the dominance relation is reversed when $q = 2.378$.

6. Concluding Remarks

A new class of nonparametric estimators for the Pickands dependence function has been proposed in this papar. It is based on the Box and Cox [2] power-

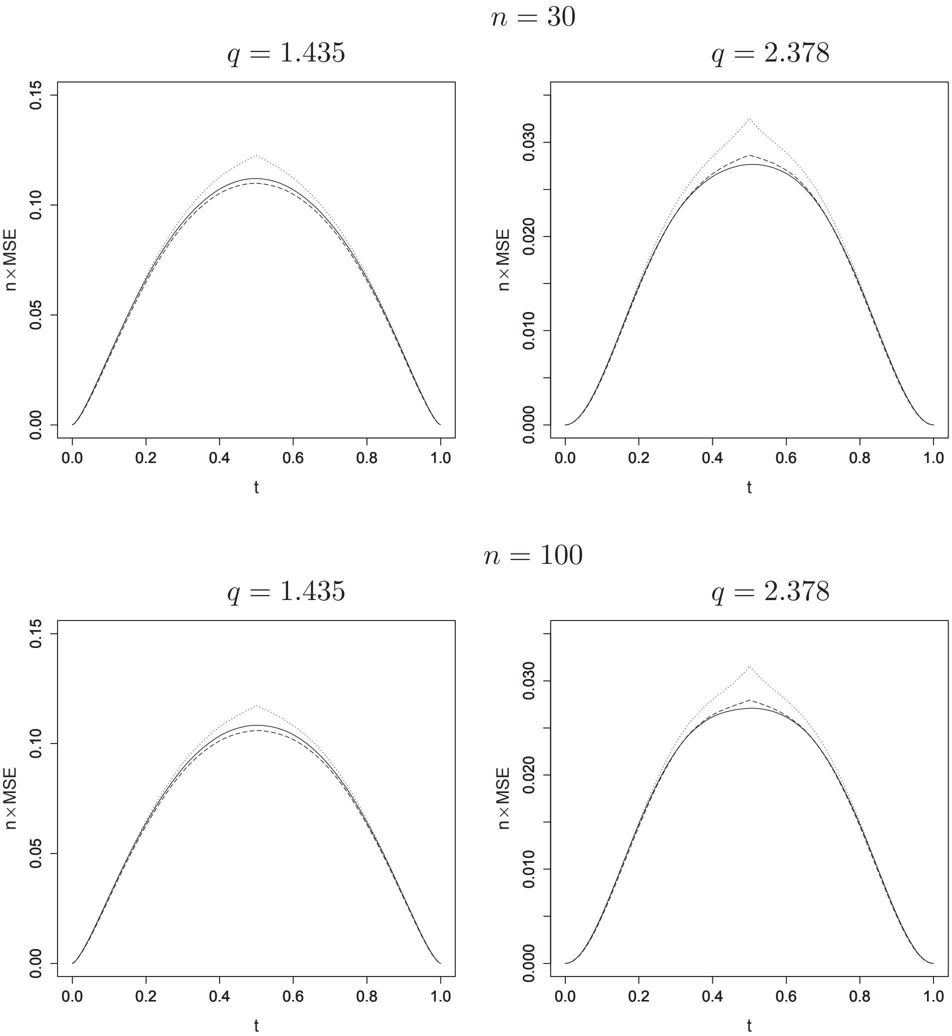

Figure 7. Pointwise mean squared error of $\hat{A}_\lambda(t;\tilde{a}_\lambda(t),\tilde{b}_\lambda(t))$ under the the logistic model (3). In each figure, the solid curve is for $\lambda=0$, the dotted curve is for $\lambda=1/3$ and the broken curve is for $\lambda=1/5$.

transformation of a key random variable $\xi(t) = \min\{X/t, Y/(1-t)\}$ which is exponentially distributed. The class contains the Pickand [19] estimator and the Capéraà, Fougères and Genest (CFG) [3] estimator. Asymptotic properties of the estimators have been investigated. Under the Marshall-Olkin model and the logistic model, both theoretical and numerical results indicate that the estimator

$$\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$$

is more preferable than the Pickand estimator and the CFG estimator. It is likely that the fifth-root transformation has a connection with a normalization of the exponential variable $\xi(t)$.

The Marshall-Olkin model and the logistic model assumed in this paper are symmetric at $t = 1/2$. It seems reasonable to suppose that performances of the estimators depend on the asymmetry of the true dependence function. It is a debatable point.

The Pickands dependence function is convex. In this paper, we have paid no attention to convexity of the estimators. A convex estimator is obtained by the projection method proposed by Fils-Villetard, Guillou and Segers [7] using $\hat{A}_{1/5}(t; \tilde{a}_{1/5}(t), \tilde{b}_{1/5}(t))$ as an initial estimator. There is room for further investigation on this point.

Appendix

Proof of Lemma 1.

Essential techniques for derivation of the covariance formula can be found in Segers [21]. For any $\lambda > 0$, some modifications are needed. For $0 < s \leq t < 1$,

$$\xi(s)\xi(t) = \min\left(\frac{X^2}{(1-s)(1-t)}, \frac{XY}{(1-s)t}, \frac{Y^2}{st}\right).$$

Thus, for $\lambda > 0$, we can express as

$$\begin{aligned}
&E\left[\{\xi(s)\xi(t)\}^\lambda\right] \\
&= \int_0^\infty \Pr\left\{\xi(s)\xi(t) > z^{1/\lambda}\right\} dz \\
&= \int_0^\infty \Pr\left\{X^2 > (1-s)(1-t)z^{1/\lambda}, XY > (1-s)tz^{1/\lambda}, Y^2 > stz^{1/\lambda}\right\} dz \\
&= \int_0^\infty E\Bigg[I\left(X^2 > (1-s)(1-t)z^{1/\lambda}\right) \\
&\qquad \times \Pr\left\{Y > \max\left(\frac{(1-s)tz^{1/\lambda}}{X}, \sqrt{stz^{1/\lambda}}\right) |X\right\}\Bigg] dz.
\end{aligned}$$

From Lemma 1 of Segers [21], conditional survival function of Y given $X=x$ can be written as

$$\Pr(Y>y|X=x)=e^{x}g(x,y),$$

where $g(x,y)=S(x,y)Q(w)$, $Q(w)=A(w)-wA'(w)$, $w=y/(x+y)$ and A' is a right-hand derivative of A.

Hence, we have

$$\begin{aligned}
&E\left[\{\xi(s)\xi(t)\}^{\lambda}\right]\\
&= \int_0^\infty E\Bigg[I\left(X^2>(1-s)(1-t)z^{1/\lambda}\right)\\
&\qquad \times e^{X}g\left(X,\max\left(\frac{(1-s)tz^{1/\lambda}}{X},\sqrt{stz^{1/\lambda}}\right)\right)\Bigg]\,dz\\
&= \int_0^\infty\int_0^\infty I\left(x^2>(1-s)(1-t)z^{1/\lambda}\right)\\
&\qquad \times g\left(x,\max\left(\frac{(1-s)tz^{1/\lambda}}{x},\sqrt{stz^{1/\lambda}}\right)\right)\,dxdz.
\end{aligned}$$

Let $(x,z)=(x,u^{\lambda}x^{2\lambda})$ and $v(u)=\max((1-s)tu,\sqrt{stu})$. Then,

$$\begin{aligned}
E\left[\{\xi(s)\xi(t)\}^{\lambda}\right] &= \int_0^\infty\int_0^\infty I((1-s)(1-t)u<1)g(x,v(u)x)\lambda u^{\lambda-1}x^{2\lambda}\,dxdz.\\
&= \int_0^{\frac{1}{(1-s)(1-t)}}\lambda u^{\lambda-1}\left\{\int_0^\infty g(x,v(u)x)x^{2\lambda}\,dx\right\}\,du.
\end{aligned}$$

Denoting $w(u)=v(u)/\{1+v(u)\}$, it can be written as

$$\begin{aligned}
&E\left[\{\xi(s)\xi(t)\}^{\lambda}\right] \qquad (20)\\
&= \int_0^{\frac{1}{(1-s)(1-t)}}\lambda u^{\lambda-1}Q(w(u))\left[\int_0^\infty x^{2\lambda}\exp\{-(1+v(u))A(w(u))x\}\,dx\right]du.\\
&= \int_0^{\frac{1}{(1-s)(1-t)}}\lambda u^{\lambda-1}Q(w(u))\frac{\Gamma(1+2\lambda)}{\{(1+v(u))A(w(u))\}^{1+2\lambda}}\,du.\\
&= \Gamma(1+2\lambda)\int_0^{\frac{s}{(1-s)^2t}}\frac{\lambda u^{\lambda-1}Q\left(\frac{\sqrt{stu}}{1+\sqrt{stu}}\right)}{\left\{(1+\sqrt{stu})A\left(\frac{\sqrt{stu}}{1+\sqrt{stu}}\right)\right\}^{1+2\lambda}}\,du\\
&\quad +\Gamma(1+2\lambda)\int_{\frac{s}{(1-s)^2t}}^{\frac{1}{(1-s)(1-t)}}\frac{\lambda u^{\lambda-1}Q\left(\frac{(1-s)tu}{1+(1-s)tu}\right)}{\left[\{1+(1-s)tu\}A\left(\frac{(1-s)tu}{1+(1-s)tu}\right)\right]^{1+2\lambda}}\,du.
\end{aligned}$$

Put $w = \sqrt{stu}/(1+\sqrt{stu})$ in the first term of (20), then

$$\begin{aligned}
\text{the first term of (20)} &= \frac{2\lambda\Gamma(1+2\lambda)}{(st)^{\lambda}} \int_0^s \frac{w^{2\lambda-1}Q(w)}{\{A(w)\}^{1+2\lambda}}\,dw \\
&= \frac{\Gamma(1+2\lambda)}{(st)^{\lambda}} \int_0^s \frac{2\lambda w^{2\lambda-1}\{A(w)-wA'(w)\}}{\{A(w)\}^{1+2\lambda}}\,dw \\
&= \frac{\Gamma(1+2\lambda)}{(st)^{\lambda}} \left[\left\{\frac{w}{A(w)}\right\}^{2\lambda}\right]_0^s = \frac{\Gamma(1+2\lambda)}{\{A(s)\}^{2\lambda}}\left(\frac{s}{t}\right)^{\lambda}.
\end{aligned}$$

Put $w = (1-s)tu/\{1+(1-s)tu\}$ in the second term of (20), then

$$\text{the second term of (20)} = \frac{\Gamma(1+2\lambda)}{(1-s)^{\lambda}t^{\lambda}} \int_s^t \frac{\lambda w^{\lambda-1}(1-w)^{\lambda}Q(w)}{\{A(w)\}^{1+2\lambda}}\,dw$$

$$\begin{aligned}
&= \frac{\Gamma(1+2\lambda)}{2(1-s)^{\lambda}t^{\lambda}} \int_s^t \left(\frac{1-w}{w}\right)^{\lambda} \left\{\left(\frac{w}{A(w)}\right)^{2\lambda}\right\}' dw \\
&= \frac{\Gamma(1+2\lambda)}{2(1-s)^{\lambda}t^{\lambda}} \left\{\left[\left(\frac{1-w}{w}\right)^{\lambda}\left(\frac{w}{A(w)}\right)^{2\lambda}\right]_s^t + \lambda\int_s^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}\,dw\right\} \\
&= \frac{\Gamma(1+2\lambda)}{2}\left[\left\{\frac{1-t}{(1-s)\{A(t)\}^2}\right\}^{\lambda} - \left\{\frac{s}{t\{A(s)\}^2}\right\}^{\lambda}\right. \\
&\qquad \left. + \frac{\lambda}{(1-s)^{\lambda}t^{\lambda}} \int_s^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}\,dw\right].
\end{aligned}$$

The result follows from

$$\begin{aligned}
\mathrm{Cov}[\varphi_{\lambda}\{\xi(s)\},\varphi_{\lambda}\{\xi(t)\}] &= \mathrm{Cov}\left[\frac{\{\xi(s)\}^{\lambda}-1}{\lambda}, \frac{\{\xi(t)\}^{\lambda}-1}{\lambda}\right] \\
&= \frac{E\left[\{\xi(s)\xi(t)\}^{\lambda}\right] - E\left[\{\xi(s)\}^{\lambda}\right]E\left[\{\xi(t)\}^{\lambda}\right]}{\lambda^2}. \qquad \square
\end{aligned}$$

Proof of Corollary 2.

Let, for $\lambda > 0$,

$$g(\lambda) = \lambda\int_0^1 u^{\lambda-1}(1-u)^{\lambda-1}\,du = \frac{2\{\Gamma(1+\lambda)\}^2}{\Gamma(1+2\lambda)}.$$

Then, by using L'Hôpital's rule, we can easily verify that, for any fixed $x > 0$,

$$\lim_{\lambda\downarrow 0} g(\lambda) = 2, \quad \lim_{\lambda\downarrow 0}\frac{2-g(\lambda)}{\lambda^2} = \frac{\pi^2}{3}, \tag{21}$$

$$\lim_{\lambda\downarrow 0}\frac{\varphi_\lambda(x)-\log x}{\lambda}=\frac{(\log x)^2}{2},\ \lim_{\lambda\downarrow 0}\frac{\varphi_\lambda(x^2)-g(\lambda)\varphi_\lambda(x)}{\lambda}=(\log x)^2. \tag{22}$$

From Lemma 1, we can express as

$$\begin{aligned}
&\frac{g(\lambda)}{\{\Gamma(1+\lambda)\}^2}\mathrm{Cov}[\varphi_\lambda\{\xi(s)\},\varphi_\lambda\{\xi(t)\}] \qquad (23)\\
=\ &\frac{2}{\lambda^2}+\lambda^{-1}\varphi_\lambda\left(\frac{1-t}{(1-s)\{A(t)\}^2}\right)+\lambda^{-1}\varphi_\lambda\left(\frac{s}{t\{A(s)\}^2}\right)\\
&+\lambda^{-1}\varphi_\lambda\left(\frac{1}{t(1-s)}\right)\int_s^t\left\{1+\lambda\varphi_\lambda\left(\frac{w(1-w)}{\{A(w)\}^2}\right)\right\}\frac{dw}{w(1-w)}\\
&-\frac{g(\lambda)}{\lambda^2}\left\{1+\lambda\varphi_\lambda\left(\frac{1}{A(s)A(t)}\right)\right\}\\
=\ &\frac{2-g(\lambda)}{\lambda^2}\\
&+\frac{1}{\lambda}\left[\varphi_\lambda\left(\frac{1-t}{(1-s)\{A(t)\}^2}\right)+\varphi_\lambda\left(\frac{s}{t\{A(s)\}^2}\right)+\int_s^t\frac{dw}{w(1-w)}\right.\\
&\qquad\left.-g(\lambda)\varphi_\lambda\left(\frac{1}{A(s)A(t)}\right)\right]\\
&+\varphi_\lambda\left(\frac{1}{t(1-s)}\right)\int_s^t\frac{dw}{w(1-w)}+\int_s^t\varphi_\lambda\left(\frac{w(1-w)}{\{A(w)\}^2}\right)\frac{dw}{w(1-w)}\\
&+\lambda\varphi_\lambda\left(\frac{1}{t(1-s)}\right)\int_s^t\varphi_\lambda\left(\frac{w(1-w)}{\{A(w)\}^2}\right)\frac{dw}{w(1-w)}.
\end{aligned}$$

On the left-hand side of (23), $g(\lambda)/\{\Gamma(1+\lambda)\}^2\to 2$ as $\lambda\downarrow 0$.

On the other hand, from (21), the first term on the right-hand side of (23) converges to $\pi^2/3$ as $\lambda\downarrow 0$. The last term of (23) converges to zero. The third term converges to

$$\begin{aligned}
&\log\frac{1}{t(1-s)}\log\frac{t(1-s)}{s(1-t)}+\int_s^t\log\left\{\frac{w(1-w)}{(A(w))^2}\right\}\frac{dw}{w(1-w)}\\
=&-\frac{1}{2}(\log t)^2-\frac{1}{2}(\log s)^2-\frac{1}{2}\{\log(1-s)\}^2-\frac{1}{2}\{\log(1-t)\}^2\\
&+2(\log t)\{\log(1-t)\}+(\log s)(\log t)+\{\log(1-s)\}\{\log(1-t)\}\\
&-2\{\log(1-s)\}(\log t)+2\int_s^t\frac{\log w}{1-w}\,dw-2\int_s^t\frac{\log A(w)}{w(1-w)}\,dw.
\end{aligned}$$

The second term can be written as

$$
\begin{aligned}
&\text{2nd term of right-hand side of (23)}\\
&= \frac{1}{\lambda}\left[\varphi_\lambda\left(\frac{1-t}{(1-s)\{A(t)\}^2}\right)+\varphi_\lambda\left(\frac{s}{t\{A(s)\}^2}\right)+\log\frac{t(1-s)}{s(1-t)}\right.\\
&\quad\left.-g(\lambda)\varphi_\lambda\left(\frac{1}{A(s)A(t)}\right)\right]\\
&= \frac{1}{\lambda}\left[\varphi_\lambda\left(\frac{1-t}{(1-s)\{A(t)\}^2}\right)-\log\left(\frac{1-t}{(1-s)\{A(t)\}^2}\right)\right]\\
&\quad+\frac{1}{\lambda}\left[\varphi_\lambda\left(\frac{s}{t\{A(s)\}^2}\right)-\log\left(\frac{s}{t\{A(s)\}^2}\right)\right]\\
&\quad-\frac{1}{\lambda}\left[\varphi_\lambda\left(\frac{1}{\{A(s)A(t)\}^2}\right)-\log\left(\frac{1}{\{A(s)A(t)\}^2}\right)\right]\\
&\quad+\frac{1}{\lambda}\left[\varphi_\lambda\left(\frac{1}{\{A(s)A(t)\}^2}\right)-g(\lambda)\varphi_\lambda\left(\frac{1}{A(s)A(t)}\right)\right].
\end{aligned}
$$

From (22), it is seen that the second term converges to

$$
\begin{aligned}
&\frac{1}{2}\left[\log\frac{1-t}{(1-s)\{A(t)\}^2}\right]^2+\frac{1}{2}\left[\log\frac{s}{t\{A(s)\}^2}\right]^2\\
&-\frac{1}{2}\left[\log\frac{1}{\{A(s)A(t)\}^2}\right]^2+\left[\log\frac{1}{A(s)A(t)}\right]^2\\
&\quad=\frac{1}{2}\{\log(1-t)\}^2+\frac{1}{2}\{\log(1-s)\}^2-\{\log(1-s)\}\{\log(1-t)\}\\
&\quad\quad+\frac{1}{2}(\log s)^2+\frac{1}{2}(\log t)^2-(\log s)(\log t)-2\left(\log\frac{1-t}{1-s}\right)\{\log A(t)\}\\
&\quad\quad-2\left(\log\frac{s}{t}\right)\{\log A(s)\}+\left\{\log\frac{A(s)}{A(t)}\right\}^2. \square
\end{aligned}
$$

Proof of Theorem 4.

Let, for $i = 1, 2, \ldots, n$,

$$W_{\lambda,i}(t) = \frac{1}{\Gamma(1+\lambda)}\left[\varphi_\lambda\{\xi_i(t)\} - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i)\right] - c_\lambda\{1-a(t)-b(t)\},$$

then, from (9),

$$\varphi_\lambda\{1/\hat{A}_\lambda(t; a(t), b(t))\} = \frac{1}{n}\sum_{i=1}^{n} W_{\lambda,i}(t).$$

It is nothing but an arithmetic average of the i.i.d. random variables $W_{\lambda,i}(t)$, $i = 1, 2, \ldots, n$, with mean

$$E\left[W_{\lambda,i}(t)\right] = \varphi_\lambda\{1/A(t)\}$$

and variance

$$\mathrm{Var}\left[W_{\lambda,i}(t)\right] = \tau_\lambda^2(t; a(t), b(t)).$$

For $\lambda \geq 0$, define a function $h_\lambda(x)$ on $[0, \infty)$ by

$$h_\lambda(x) = \begin{cases} (1+\lambda x)^{-1/\lambda}, & \lambda > 0 \\ e^{-x}, & \lambda = 0. \end{cases}$$

Then, the k-th derivative of $h_\lambda(x)$ is given by

$$h_\lambda^{(k)}(x) = (-1)^k (1+\lambda x)^{-k} h_\lambda(x) \prod_{j=1}^{k-1} \{1 + j\lambda\},$$

and which is bounded on $[0, \infty)$ for $k = 1, 2, \ldots,$. Noting

$$\hat{A}_\lambda(t; a(t), b(t)) = h_\lambda\left(\frac{1}{n}\sum_{i=1}^{n} W_{\lambda,i}(t)\right) \quad \text{and} \quad A(t) = h_\lambda\left(\varphi_\lambda\{1/A(t)\}\right),$$

from Theorem 4.2.1 of Lehmann [16], we obtain

$$\begin{aligned} E\left[\hat{A}_\lambda(t; a(t), b(t))\right] &= A(t) + \frac{\tau_\lambda^2(t; a(t), b(t))}{2n} h_\lambda^{(2)}(\varphi_\lambda\{1/A(t)\}) + O\left(\frac{1}{n^2}\right), \\ \mathrm{Var}\left[\hat{A}_\lambda(t; a(t), b(t))\right] &= \frac{\tau_\lambda^2(t; a(t), b(t))}{n} \left\{h_\lambda^{(1)}(\varphi_\lambda\{1/A(t)\})\right\}^2 + O\left(\frac{1}{n^2}\right) \end{aligned}$$

and

$$\begin{aligned} &\sqrt{n}\left\{\hat{A}_\lambda(t; a(t), b(t)) - A(t)\right\} \\ &\quad \xrightarrow{L} N\left(0, \tau_\lambda^2(t; a(t), b(t)) \left\{h_\lambda^{(1)}(\varphi_\lambda\{1/A(t)\})\right\}^2\right). \ \square \end{aligned}$$

Proof of equation (13)

We present a proof for $s < 1/2 \leq t$. Equations for other cases can be shown similarly.

$$\begin{aligned} &\lambda \int_s^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}\, dw \\ &\quad = \lambda \int_s^{\frac{1}{2}} \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{(1-\theta w)^{2\lambda}}\, dw + \lambda \int_{1-t}^{\frac{1}{2}} \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{(1-\theta w)^{2\lambda}}\, dw. \end{aligned}$$

Changing variable by $y = (1-\theta)w/(1-\theta w)$, we have

$$\begin{aligned}
&\lambda \int_s^t \frac{w^{\lambda-1}(1-w)^{\lambda-1}}{\{A(w)\}^{2\lambda}}\, dw \\
&= (1-\theta)^{-\lambda} \left\{ \lambda \int_{\frac{1-\theta}{2-\theta}}^{\frac{(1-\theta)s}{1-\theta s}} y^{\lambda-1}(1-y)^{\lambda-1}\, dw \right. \\
&\qquad \left. + \lambda \int_{\frac{1-\theta}{2-\theta}}^{\frac{(1-\theta)(1-t)}{1-\theta(1-t)}} y^{\lambda-1}(1-y)^{\lambda-1}\, dw \right\} \\
&= (1-\theta)^{-\lambda} \left\{ 2\left(\frac{1-\theta}{2-\theta}\right)^{\lambda} f_\lambda\left(\frac{1-\theta}{2-\theta}\right) - \left(\frac{(1-\theta)s}{1-\theta s}\right)^{\lambda} f_\lambda\left(\frac{(1-\theta)s}{1-\theta s}\right) \right. \\
&\qquad \left. - \left(\frac{(1-\theta)(1-t)}{1-\theta(1-t)}\right)^{\lambda} f_\lambda\left(\frac{(1-\theta)(1-t)}{1-\theta(1-t)}\right) \right\}. \square
\end{aligned}$$

Proof of the equation (17)

Under complete dependence, $X_i = Y_i$, $i = 1, 2, \ldots, n$ with probability one. In this case, we have

$$\xi_i(t) = \min\left(\frac{X_i}{1-t}, \frac{Y_i}{t}\right) = \frac{X_i}{\max(1-t, t)} = \frac{X_i}{A(t)},$$

and

$$\begin{aligned}
&\varphi_\lambda(\xi_i(t)) - a(t)\varphi_\lambda(X_i) - b(t)\varphi_\lambda(Y_i) \\
&= \lambda^{-1}\left[X^\lambda\left\{(A(t))^{-\lambda} - a(t) - b(t)\right\} - 1 + a(t) + b(t)\right] \\
&= \lambda^{-1}\{a(t) + b(t) - 1\}.
\end{aligned}$$

Thus, from (9), it holds that

$$\begin{aligned}
&\varphi_\lambda(1/\hat{A}_\lambda(t; a(t), b(t))) \\
&= \frac{1}{n\Gamma(1+\lambda)} \sum_{i=1}^{n} \lambda^{-1}\{a(t) + b(t) - 1\} - c_\lambda\{1 - a(t) - b(t)\} \\
&= \frac{a(t) + b(t) - 1}{\lambda} = \frac{(A(t))^{-\lambda} - 1}{\lambda} = \varphi_\lambda(1/A(t)). \square
\end{aligned}$$

References

[1] Beirlant, J., Goegebeur, Y. and Segers, J. and Teugels, J. (2004). *Statistics of Extremes, Theory and Applications*, Wiley, Chichester.

[2] Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society, Series B*, **26**, (1964), 211-252.

[3] Capéraà, P., Fougères, A. L. and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas, *Biometrika*, **84**, 567-577.

[4] 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 York.

[5] Deheuvels, P. (1991). On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions, *Statistics & Probability Letters*, **12**, 429-439.

[6] Deheuvels, P. and Tiago de Oliveira, J. (1989). On the non-parametric estimation of the bivariate extreme-value distributions, *Statistics & Probability Letters*, **8**, 315-323.

[7] Fils-Villetard, A., Guillou, A., and Segers, J. (2008). Projection estimators of Pickands dependence functions, *The Canadian Journal of Statistics*, **36**, 369-382.

[8] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New York.

[9] Gumbel, E. J. (1960). Bivariate exponential distributions, *Journal of the American Statistical Association*, **55**, 698-707.

[10] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory, An Introduction*, Springer, New York.

[11] Hall, P. and Tajvidi, N. (2000). Distribution and dependence-function estimation for bivariate extreme-value distributions, *Bernoulli*, **6**, 835-844.

[12] Jiménz, J. R., Villa-Diharce, E. and Flores, M. (2001). Nonparametric estimation of dependence function in bivariate extreme value distributions, *Journal of Multivariate Analysis*, bf 76, 159-191.

[13] Joe, H. (1997). *Multivariate Models and Dependence Concept*, Chapman and Hall, London.

[14] Johnson, N. L., Kemp, A. W. and Kotz, S. (2005). *Univariate Discrete Distributions, 3rd ed.*, Wiley, New York.

[15] Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions*, Imperial College Press, London.

[16] Lehmann, E. L. (1999). *Elements of Large-Sample Theory*, Springer, New York.

[17] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution, *Journal of the American Statistical Association*, **62**, 30-44.

[18] Nelsen, R. B. (2006). *An Introduction to Copulas, 2nd ed.*, Springer, Berlin.

[19] Pickands, J. (1981). Multivariate extreme value distributions, *Proceedings of the 43rd Session of the ISI*, Buenos Aires, **49**, 859-878.

[20] Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*, Springer, New York.

[21] Segers, J. (2008). Non-parametric inference for bivariate extreme-value copulas, In *Topics in Extreme Values* (M. Ahsanullah and S. N. U. A. Kirmani, eds.), 185-207. Nova Science Publishers, New York.

[22] Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation, *Biometrika*, **75**, 3, 397-415.

[23] Tiago de Oliveira, J. (1989). Intrinsic estimation of the dependence structure for bivariate extremes, *Statistics & Probability Letters*, **8**, 213-218.

Estimators of Bivariate Extreme Value Copulas

SUZUKAWA Akio

Abstract

Extreme value copulas are the limiting copulas of component-wise maxima. A bivariate extreme value copulas can be represented by a convex function called Pickands dependence function. In this paper we consider nonparametric estimation of the Pickands dependence function. Several estimators have been proposed. They can be classified into two types: Pickands-type estimators and Capéraà-Fougères-Genest-type estimators. We propose a new class of estimators, which contains these two types of estimators. Asymptotic properties of the estimators are investigated, and asymptotic efficiencies of them are discussed under Marshall-Olkin copulas.

Keywords

bivariate exponential distribution, extreme value distribution, Pickands dependence function