Nonparametric estimation of the dependence function for a multivariate extreme value distribution

Dabao Zhang\textsuperscript{a}, Martin T. Wells\textsuperscript{b,c,*}, Liang Peng\textsuperscript{d}

\textsuperscript{a}Department of Statistics, Purdue University, West Lafayette, IN 47907, USA
\textsuperscript{b}Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA
\textsuperscript{c}Department of Statistical Science, Cornell University, Ithaca, NY 14853, USA
\textsuperscript{d}School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA

Received 5 June 2006
Available online 19 December 2006

Abstract

Understanding and modeling dependence structures for multivariate extreme values are of interest in a number of application areas. One of the well-known approaches is to investigate the Pickands dependence function. In the bivariate setting, there exist several estimators for estimating the Pickands dependence function which assume known marginal distributions [J. Pickands, Multivariate extreme value distributions, Bull. Internat. Statist. Inst., 49 (1981) 859–878; P. Deheuvels, On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions, Statist. Probab. Lett. 12 (1991) 429–439; P. Hall, N. Tajvidi, Distribution and dependence-function estimation for bivariate extreme-value distributions, Bernoulli 6 (2000) 835–844; P. Capéràà, A.-L. Fougeres, C. Genest, A nonparametric estimation procedure for bivariate extreme value copulas, Biometrika 84 (1997) 567–577]. In this paper, we generalize the bivariate results to \( p \)-variate multivariate extreme value distributions with \( p \geq 2 \). We demonstrate that the proposed estimators are consistent and asymptotically normal as well as have excellent small sample behavior.

© 2006 Elsevier Inc. All rights reserved.

AMS 1991 subject classification: 62G32; 62H20

Keywords: Copulas; Dependence function; Empirical distribution; Gaussian process; Multivariate extreme value distribution

* Corresponding author.

E-mail addresses: zhangdb@stat.purdue.edu (D. Zhang), mtw1@cornell.edu (M.T. Wells), peng@math.gatech.edu (L. Peng).

0047-259X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
1. Introduction

Suppose \((X_{i1}, \ldots, X_{ip})\) for \(i = 1, \ldots, n\) are iid random vectors with a multivariate extreme-value distribution \(F\) (as defined in [13]) and let \(F_j(x)\) denote the marginal distribution of \(X_{ij}\) for \(j = 1, \ldots, p\). Modeling the dependence structure of \(F\) is of interest in a number of contexts such as environmental resources management, financial risk management, and data network analysis. When \(p = 2\), Pickands [12] constructed a copula-type representation

\[
C(u, v) = P(F_1(X_{i1}) \leq u, F_2(X_{i2}) \leq v) = \exp \left\{ \log(uv) A \left( \frac{\log(u)}{\log(uv)} \right) \right\},
\]

for \(0 \leq u, v \leq 1\), where \(A(t)\), called dependence function, is a convex function defined on \([0, 1]\) and satisfies \(\max(t, 1 - t) \leq A(t) \leq 1\) for all \(0 \leq t \leq 1\). It is known that the dependence function \(A(t)\) is related to some well-known nonparametric dependence measures such as Kendall’s and Spearman’s measures of dependence; see Tawn [17] and Ghoudi et al. [8]. For more on nonparametric estimation of the dependence function \(A(t)\), see Pickands [12], Deheuvels [4], Tiago de Oliveira [18], Capéraà et al. [1], and Hall and Tajvidi [10].

Recently, Falk and Reiss [7] generalized the above dependence structure to the case \(p > 2\). That is,

\[
C(u_1, \ldots, u_p) := P(F_1(X_{i1}) \leq u_1, \ldots, F_p(X_{ip}) \leq u_p)
= \exp \left\{ (\sum_{j=1}^p \log u_j) A \left( \frac{\sum_{j=1}^p \log u_j}{\sum_{j=1}^p \log u_j} \right) \right\},
\]

Some characterization properties of the dependence function \(A(s_1, \ldots, s_{p-1})\) can be found in Falk and Reiss [7].

Although the dependence structure of \(F\) can be described by other types of \((p - 1)\)-dimensional dependence function ([9, 13]), we focus on estimating the Pickands type of dependence function \(A(\cdot)\) defined above. Let \(Y_{ij} = -\log\{F_j(X_{ij})\}, j = 1, \ldots, p\), then the \(Y_{ij}\) are distributed as the standard exponential distribution. The estimators by Pickands [12], Deheuvels [4], and Hall and Tajvidi [10] are directly generalized as, respectively,

\[
\hat{A}^P(s_1, \ldots, s_{p-1}) = \frac{n}{\sum_{i=1}^n \bigwedge_{j=1}^p Y_{ij}/s_j},
\]

\[
\hat{A}^D(s_1, \ldots, s_{p-1}) = \frac{n}{\sum_{i=1}^n \bigwedge_{j=1}^p Y_{ij}/s_j - n \sum_{j=1}^p s_j \bar{Y}_j + n},
\]

\[
\hat{A}^{HT}(s_1, \ldots, s_{p-1}) = \frac{n}{\sum_{i=1}^n \bigwedge_{j=1}^p Y_{ij}/s_j},
\]

where \(\bar{Y}_j = \sum_{i=1}^n Y_{ij}/n, 1 \leq j \leq p\) and \(\sum_{j=1}^p s_j = 1\). As in the case \(p = 2\) studied by Hall and Tajvidi [10], one may expect that both \(\hat{A}^D(s_1, \ldots, s_{p-1})\) and \(\hat{A}^{HT}(s_1, \ldots, s_{p-1})\) improve the small sample performance over \(\hat{A}^P(s_1, \ldots, s_{p-1})\), particularly when \((s_1, \ldots, s_{p-1})\) is close to the boundary of the simplex

\[
S_{p-1} := \{(s_1, \ldots, s_{p-1}) : \sum_{j=1}^{p-1} s_j \leq 1 \text{ and } s_j \geq 0 \text{ for } j = 1, \ldots, p - 1\}.
\]
However, the nonparametric estimator for $A(\cdot)$ proposed by Capéraà et al. [1] (hereafter CFG estimator) performs much better than the Pickands estimator and its variants when $p = 2$. In this paper we generalize the CFG estimator to the case $p > 2$. As a matter of fact, this generalization was posted as an open question by Capéraà et al. [1]. We note that the setup in this paper is quite different from another extensive study of nonparametric estimation of the dependence function and spectral measure, where $F$ is assumed to be in the domain of attraction of a bivariate extreme value distribution rather than an exact bivariate extreme value distribution; see Huang [11], Einmahl et al. [6,5].

The paper is organized as follows. In Section 2, we generalize the CFG estimator to dimension larger than two and state the large sample properties. In Section 3, a simulation study is conducted to compare these estimators. A data analysis of the dependence structures among the water levels observed in four stations around Lake Ontario is given in Section 4. All of the proofs are put into the Appendix.

2. The proposed estimator and the main results

For $(s_1, \ldots, s_{p-1}) \in S_{p-1}$. Note that $A(s_1, \ldots, s_{p-1}) = 1$ whenever there exists $j$ such that $s_j = 1$. We will assume $s_j < 1$, $j = 1, 2, \ldots, p$, in the following discussion. And, for any function of $(s_1, \ldots, s_{p-1})$, its value at $s_j = 0$ is defined by its limit as $s_j \downarrow 0$.

As in the construction of the CFG estimator, we define a new set of auxiliary random variables

$$Z_{ij} = \frac{\sqrt{l \neq j \log \{F_i(X_{il})\} / s_i}}{\log \{F_j(X_{ij})\} / s_j} + \frac{\sqrt{l \neq j \log \{F_i(X_{il})\} / s_i}}{\log \{F_j(X_{ij})\} / s_j}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, p,$$  

(2.1)

where $(s_1, \ldots, s_{p-1}) \in S_{p-1}$ and $s_p = 1 - \sum_{j=1}^{p-1} s_j$. Then, for $j = 1, \ldots, p-1$,

$$H_j(z) := P(Z_{ij} \leq z) = z + z(1-z)$$

$$\times \frac{\hat{\partial}}{\partial z} \log A \left( \frac{zs_1}{1-s_j}, \ldots, \frac{zs_{j-1}}{1-s_j}, 1-z, \frac{zs_{j+1}}{1-s_j}, \ldots, \frac{zs_{p-1}}{1-s_j} \right)$$  

(2.2)

and

$$H_p(z) := P(Z_{ip} \leq z) = z + z(1-z) \times \frac{\hat{\partial}}{\partial z} \log A \left( \frac{zs_1}{1-s_p}, \ldots, \frac{zs_{p-1}}{1-s_p} \right)$$  

(2.3)

(see the Appendix for the proof). Hence solving the differential equation we have

$$\log A(s_1, \ldots, s_{p-1}) = \int_0^{1-s_j} \frac{H_j(z) - z}{z(1-z)} dz$$

for $j = 1, \ldots, p$. Based on the representations in (2.1)–(2.3), we can construct an estimate of $A(s_1, \ldots, s_{p-1})$ by

$$\hat{A}_j(s_1, \ldots, s_{p-1}) = \exp \left\{ \int_0^{1-s_j} \frac{\hat{H}_j(z) - z}{z(1-z)} dz \right\},$$  

(2.4)
where
\[
\hat{H}_j(z) = \frac{1}{n} \sum_{i=1}^{n} I(Z_{ij} \leq z), \quad j = 1, \ldots, p.
\]

As in Capéraà et al. [1], we propose the following weighted estimator:
\[
\hat{A}(s_1, \ldots, s_{p-1}) = \prod_{j=1}^{p} \left( \hat{A}_j(s_1, \ldots, s_{p-1}) \right)^{\hat{\lambda}_j(s_1, \ldots, s_{p-1})},
\]
where \(\hat{\lambda}_1(s_1, \ldots, s_{p-1}), \ldots, \hat{\lambda}_p(s_1, \ldots, s_{p-1})\) are nonnegative weight functions and satisfy
\[
\sum_{j=1}^{p} \hat{\lambda}_j(s_1, \ldots, s_{p-1}) = 1.
\]

We remark that \(\hat{A}(s_1, \ldots, s_{p-1})\) is a discontinuous function since \(\hat{H}_j(z)\) depends on \(s_1, \ldots, s_{p-1}\). Hence, it is not a convex function as well. By defining \(\hat{A}_j(s_1, \ldots, s_{p-1}) = 1\) whenever \(s_j = 1\), it is easy to check that \(\lim_{s_j \uparrow 1} \hat{A}(s_1, \ldots, s_{p-1}) = 1\) if \(\lim_{s_j \uparrow 1} \hat{\lambda}_l(s_1, \ldots, s_{p-1}) \log s_l = 0\) for all \(l \neq j\). A simple choice of \(\hat{\lambda}_j(s_1, \ldots, s_{p-1})\) is to set \(\hat{\lambda}_j(s_1, \ldots, s_{p-1}) = s_j\) for \(j = 1, \ldots, p\).

In Remark 3 below we provide a theoretical optimal choice of \(\hat{\lambda}_j(s_1, \ldots, s_{p-1})\), which involves solving \(p + 1\) simultaneous equations.

Now for any fixed \((s_1, \ldots, s_{p-1}) \in S_{p-1}\) and \(s_p\), let \(B(u_1, \ldots, u_p)\) denote a Gaussian process on \([0, 1]^p\) with \(E\{B(u_1, \ldots, u_p)\} = 0\) and
\[
E\{B(u_1, \ldots, u_p)B(v_1, \ldots, v_p)\} = H(u_1 \wedge v_1, \ldots, u_p \wedge v_p)

- H(u_1, \ldots, u_p)H(v_1, \ldots, v_p)
\]
for \((u_1, \ldots, u_p) \in [0, 1]^p\) and \((v_1, \ldots, v_p) \in [0, 1]^p\), where \(H(u_1, \ldots, u_p) = P(Z_{11} \leq u_1, \ldots, Z_{1p} \leq u_p)\). Note that both \(B(u_1, \ldots, u_p)\) and \(H(u_1, \ldots, u_p)\) depend on \((s_1, \ldots, s_{p-1})\).

Let \(B_j(u)\) denote \(B(u_1, \ldots, u_p)\) with \(u_j = u\) and \(u_i = 1\) for all \(i \neq j\).

Our main theoretical results are that the proposed estimator in (2.5) is uniformly consistent and asymptotically normal. The proof of the result is given in the Appendix.

**Theorem 1.** Suppose \(A(s_1, \ldots, s_{p-1})\) has bounded first partial derivatives. For any fixed \((s_1, \ldots, s_{p-1}) \in S_{p-1}\), we have
\[
\sup_{(s_1, \ldots, s_{p-1}) \in S_{p-1}} |\hat{A}(s_1, \ldots, s_{p-1}) - A(s_1, \ldots, s_{p-1})| \xrightarrow{p} 0
\]
and
\[
\sqrt{n}\{\log \hat{A}(s_1, \ldots, s_{p-1}) - \log A(s_1, \ldots, s_{p-1})\}

\xrightarrow{d} \sum_{j=1}^{p} \lambda_j(s_1, \ldots, s_{p-1}) \int_{0}^{1-s_j} \frac{B_j(z)}{z(1-z)} \, dz

\xrightarrow{d} N(0, \sigma^2),
\]
(2.7)
where
\[
\sigma^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} \hat{\lambda}_i(s_1, \ldots, s_{p-1}) \hat{\lambda}_j(s_1, \ldots, s_{p-1}) \times \int_{0}^{1-s_i} \int_{0}^{1-s_j} \frac{H_{ij}(z_1, z_2) - H_i(z_1)H_j(z_2)}{z_1z_2(1-z_1)(1-z_2)} \, dz_2 \, dz_1
\] (2.8)

and

\[
H_{ij}(z_1, z_2) = P(Z_{1i} \leq z_1, Z_{1j} \leq z_2).
\]

**Remark 1.** Although (2.7) is only a point-wise convergence result, not in \(D([0, 1]^{p-1})\), it is obvious that the weak convergence result is true when \(p = 2\). As far as we know, no asymptotic normality result exists in the literature for estimating the Pickands dependence function by assuming marginal distribution unknown. Consequently, it would be interesting to derive the large sample properties for \(\hat{A}(s_1, \ldots, s_{p-1})\) with \(F_j(x)\) replaced by a parametric or nonparametric estimate.

**Remark 2.** A simple consistent estimator for \(\sigma^2\) is to replace \(H_i(z_1)\), \(H_j(z_2)\) and \(H_{ij}(z_1, z_2)\) in (2.8) by \(\hat{H}_i(z_1)\), \(\hat{H}_j(z_2)\) and \(\hat{H}_{ij}(z_1, z_2)\), respectively.

**Remark 3.** An optimal choice of weights \(\hat{\lambda}_i(s_1, \ldots, s_{p-1})\) can be obtained by minimizing \(\sigma^2\). By the standard argument of Lagrange multiplier, the optimization problem reduces to solving the following equations

\[
\begin{align*}
\sum_{i=1}^{p} \hat{\lambda}_i(s_1, \ldots, s_{p-1}) \int_{0}^{1-s_i} \int_{0}^{1-s_j} \frac{H_{ij}(z_1, z_2) - H_i(z_1)H_j(z_2)}{z_1z_2(1-z_1)(1-z_2)} \, dz_2 \, dz_1 = \lambda, \\
\sum_{i=1}^{p} \hat{\lambda}_i(s_1, \ldots, s_p) = 1,
\end{align*}
\]

with respect to \(\hat{\lambda}_1(s_1, \ldots, s_{p-1}), \ldots, \hat{\lambda}_p(s_1, \ldots, s_{p-1})\) and \(\lambda\). Therefore, an estimated optimal choice can be obtained by solving the above equations with \(H_i\) and \(H_{ij}\) replaced by \(\hat{H}_i\) and \(\hat{H}_{ij}\), respectively.

3. Simulation study

To evaluate the performance of different estimators for the Pickands dependence function, we used the algorithms in Stephenson [15] to simulate trivariate extreme values whose dependence functions are logistic-type [17]. Explicitly, the logistic dependence function of trivariate extreme values is

\[
A(s_1, s_2) = \{\theta' s_1^r + \phi' s_2^r\}^{1/r} + \{\theta' s_2^r + \phi' s_3^r\}^{1/r} + \{\theta' s_3^r + \phi' s_1^r\}^{1/r}
\]

\[+ \psi\{s_1^r + s_2^r + s_3^r\}^{1/r} + 1 - \theta - \phi - \psi,
\]

where \(s_3 \equiv 1 - s_1 - s_2\), \(r \geq 1\) and \(0 \leq \theta, \phi, \psi \leq 1\). Here, we consider the symmetric logistic dependence function with \(r = 3\), \(\theta = \phi = 0\), \(\psi = 1\), and the asymmetric logistic dependence
Table 1

Mean integrated square errors and maximum mean square errors in simulation study

<table>
<thead>
<tr>
<th></th>
<th>n = 25</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLDF</td>
<td>ALDF</td>
<td>SLDF</td>
<td>ALDF</td>
<td>SLDF</td>
</tr>
<tr>
<td>MISE \times 10^5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{A}^p)</td>
<td>1117</td>
<td>1500</td>
<td>487</td>
<td>935</td>
</tr>
<tr>
<td>(\hat{A}^d)</td>
<td>143</td>
<td>493</td>
<td>70</td>
<td>273</td>
</tr>
<tr>
<td>(\hat{A}^{HT})</td>
<td>25</td>
<td>360</td>
<td>13</td>
<td>187</td>
</tr>
<tr>
<td>(\hat{A})</td>
<td>27</td>
<td>234</td>
<td>13</td>
<td>127</td>
</tr>
<tr>
<td>MMSE \times 10^5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{A}^p)</td>
<td>4948</td>
<td>4250</td>
<td>2206</td>
<td>2427</td>
</tr>
<tr>
<td>(\hat{A}^d)</td>
<td>346</td>
<td>1357</td>
<td>169</td>
<td>663</td>
</tr>
<tr>
<td>(\hat{A}^{HT})</td>
<td>114</td>
<td>1120</td>
<td>56</td>
<td>555</td>
</tr>
<tr>
<td>(\hat{A})</td>
<td>64</td>
<td>719</td>
<td>32</td>
<td>356</td>
</tr>
</tbody>
</table>

The results are for the symmetric logistic dependence function (SLDF) with \((r = 3, \theta = \phi = 0, \psi = 1)\) and the asymmetric logistic dependence function (ALDF) with \((r = 6, \theta = 0.6, \phi = 0.3, \psi = 0)\).

...
Fig. 1. Biases of $\hat{\mathbf{A}}^P$ (solid lines), $\hat{\mathbf{A}}^D$ (dashed lines), $\hat{\mathbf{A}}^{HT}$ (dash-dot lines) and $\hat{\mathbf{A}}$ (dotted lines) for trivariate extreme values along the line $s_1 = s_2 = s$. 
Fig. 2. Mean square errors of $\hat{A}^P$ (solid lines), $\hat{A}^D$ (dashed lines), $\hat{A}^{HT}$ (dash-dot lines) and $\hat{A}$ (dotted lines) for trivariate extreme values along the line $s_1 = s_2 = s$. 
4. An application to the water level data

Water levels of the Great Lakes are monitored at different intake stations (see http://glakesonline.nos.noaa.gov). Here we explore the dependence structures among the water levels observed in four stations around Lake Ontario, i.e., Cape Vincent, Niagara Intake, Oswego and Rochester. The annual maximum water levels and annual minimum water levels were recorded from 1963 to 2005, respectively. Using the ismев package [2,16], we take maximum-likelihood fitting for the generalized extreme value distributions for either maximum or minimum water levels at each station. All observed extreme values are parametrically transformed to have standard exponential marginal distributions. We apply the new approach with weights \( \lambda_j(s_1, \ldots, s_{p-1}) = s_j \) to estimate the dependence functions for all trivariate extreme values. Some of the estimated dependence functions are shown in Fig. 3.

Estimated dependence functions for bivariate extreme values can also be read from Fig. 3. For both maximum and minimum water levels, the estimated dependence function of (Oswego, Niagara, Rochester) implies that the annual extreme values at Oswego and Rochester are closely associated with each other but they are weakly associated to the annual extreme values at Niagara Intake. However, the association of the annual minimum values at Niagara Intake to those at either Oswego/Rochester is stronger than the corresponding association of the annual maximum values.

![Fig. 3. Estimated dependence functions for annual maximum values (a and c) and annual minimum values (b and d) in the water level data.](image-url)
The strong association between annual extreme values at Oswego and Rochester is also demonstrated in the estimated dependence function of (Oswego, Rochester, Cape Vincent). Indeed, Fig. 3c and d show that the annual extreme values at these three different stations have a very large chance to increase/decrease simultaneously.

**Acknowledgments**

We thank one associate editor and two reviewers for their helpful comments. Peng’s research was supported by NSF Grant DMS 04-03443 and a Humboldt Research Fellowship. Wells’s research was supported by NSF Grant DMS 02-04252.

**Appendix A.**

**Proofs of (2.2) and (2.3).** We only consider the case when \( j = 1 \) and \( s_1 \neq 1 \) since others are similar. Put \( \tilde{s}_j = s_j / (1 - s_1) \), \( j = 2, \ldots, p \). Note that

\[
G(u, v) := \Pr \left( \log \left\{ F_1(X_{11}) \right\} \leq u, \sqrt[p]{\log \left\{ F_j(X_{1j}) \right\}} / \tilde{s}_j \leq v \right) \\
= \Pr \left( F_1(X_{11}) \leq e^u, \sqrt[p]{\log \left\{ F_j(X_{1j}) \right\}} / \tilde{s}_j \leq e^v \right) \\
= \exp \left\{ (u + v) A \left( \frac{u}{u + v}, \frac{v}{u + v}, \tilde{s}_2, \ldots, \frac{v}{u + v} \tilde{s}_{p-1} \right) \right\}. \tag{A.1}
\]

Let \( G_2(u, v) = \frac{\partial}{\partial v} G(u, v) \). Then

\[
H_1(z) = \Pr \left( Z_{11} \leq z, \sqrt[p]{\log \left\{ F_j(X_{1j}) \right\}} / \tilde{s}_j \leq 0 \right) \\
= \Pr \left( \log \left\{ F_1(X_{11}) \right\} \leq \frac{1 - z}{z} \times \sqrt[p]{\log \left\{ F_j(X_{1j}) \right\}}, \sqrt[p]{\log \left\{ F_j(X_{1j}) \right\}} / \tilde{s}_j \leq 0 \right) \\
= \int_{s \leq t(1-z)/z, t \leq 0} dG(s, t) \\
= \int_{-\infty}^{0} G_2 \left( \frac{1 - z}{z}, t, t \right) \, dt.
\]

By (A.1),

\[
G_2 \left( \frac{1 - z}{z}, t, t \right) = \exp \left\{ t A(1 - z, z\tilde{s}_2, \ldots, z\tilde{s}_{p-1})/z \right\} \\
\times \left\{ A(1 - z, z\tilde{s}_2, \ldots, z\tilde{s}_{p-1}) + (1 - z) \frac{\partial}{\partial z} A(1 - z, z\tilde{s}_2, \ldots, z\tilde{s}_{p-1}) \right\}.
\]

Hence

\[
H_1(z) = z + z(1 - z) \times \frac{\partial}{\partial z} \log A(1 - z, z\tilde{s}_2, \ldots, z\tilde{s}_{p-1}). \quad \Box
\]
Proof of Theorem 1. Set
\[
\hat{H}(u_1, \ldots, u_p) = \frac{1}{n} \sum_{i=1}^{n} I(Z_{i1} \leq u_1, \ldots, Z_{ip} \leq u_p).
\]

It follows from Theorem B of Csörgő and Horváth [3] that
\[
\sup_{0 \leq u_1, \ldots, u_p \leq 1} |\sqrt{n}\{\hat{H}(u_1, \ldots, u_p) - H(u_1, \ldots, u_p)\} - B(u_1, \ldots, u_p)| = O_p(n^{-\frac{1}{2p-2}} \log(n)).
\] (A.2)

Hence, for \(1 \leq j \leq p\),
\[
\sup_{0 \leq u_j \leq 1} |\sqrt{n}\{\hat{H}_j(u_j) - H_j(u_j)\} - B_j(u_j)| = O_p(n^{-\frac{1}{2p-2}} \log(n)).
\] (A.3)

It is known that there exists \(v \in (0, \frac{1}{2})\) such that
\[
\sup_{0 \leq z \leq 1} \sqrt{n}|\hat{H}_j(z) - H_j(z)| / (H_j(z)(1 - H_j(z)))^v = O_p(1)
\] (A.4)
(see [14, Theorem 1, p. 140]).

Let
\[
D_j(z) = \frac{\partial}{\partial z} \log A \left( \frac{z s_1}{1 - s_j}, \ldots, \frac{z s_{j-1}}{1 - s_j}, \frac{1 - z}{1 - s_j}, \ldots, \frac{z s_{p-1}}{1 - s_j} \right),
\]
then \(D_j(z)\) is the derivative of \(\log A\) along a straight line within \(S_{p-1}\). Therefore, \(|D_j(z)|\) is bounded, following the conditions in Theorem 1 and the fact that \(1 \leq pA(s_1, \ldots, s_{p-1}) \leq p\) as established by Falk and Reiss [7, p. 429]. Therefore,
\[
\frac{H_j(z)(1 - H_j(z))}{z(1 - z)} = [1 + (1 - z)D_j(z)][1 - zD_j(z)]
\]
is also bounded. Note that
\[
\frac{\hat{H}_j(z) - H_j(z)}{z(1 - z)} = \frac{\hat{H}_j(z) - H_j(z)}{(H_j(z)(1 - H_j(z)))^v} \left\{ \frac{H_j(z)(1 - H_j(z))}{z(1 - z)} \right\}^v [z(1 - z)]^{v-1}.
\] (A.5)

The consistency result follows from the continuity of the integral and \(\log A\) as well as from the definition of \(\log A\) in (2.4).

Now for the proof of asymptotic normality, consider
\[
\sqrt{n} I_1 = o_p(1).
\] (A.6)
When $s_j > 0$, it follows from (A.3) that
\[
\left| \int_{1/n}^{1-s_j} \sqrt{n} \left( \hat{H}_j(z) - H_j(z) \right) \frac{1}{z(1-z)} \, dz - \int_{1/n}^{1-s_j} \frac{B_j(z)}{z(1-z)} \, dz \right| = O_p \left( \frac{1}{n^{4/p^2} \log n} \int_{1/n}^{1-s_j} \frac{1}{z(1-z)} \, dz \right) = o_p(1).
\]
Hence
\[
\sqrt{n} I_2 \xrightarrow{d} \int_0^{1-s_j} \frac{B_j(z)}{z(1-z)} \, dz. \tag{A.7}
\]
When $s_j = 0$, we can show (A.7) holds similarly by writing
\[
I_2 = \int_{1/n}^{1-1/n} \frac{\hat{H}_j(z) - H_j(z)}{z(1-z)} \, dz + \int_{1-1/n}^{1} \frac{\hat{H}_j(z) - H_j(z)}{z(1-z)} \, dz.
\]
Thus, it follows from (A.6) and (A.7) that
\[
\sqrt{n} \{ \log \hat{A}_j(s_1, \ldots, s_{p-1}) - \log A_j(s_1, \ldots, s_{p-1}) \} \xrightarrow{d} \int_0^1 \frac{B_j(z)}{z(1-z)} \, dz,
\]
which implies the second result of the theorem. \qed

References