A note on the maximum correlation for Baker’s bivariate distributions with fixed marginals

G.D. Lin\textsuperscript{a,}\textsuperscript{*}, J.S. Huang\textsuperscript{b}

\textsuperscript{a} Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan, ROC
\textsuperscript{b} Department of Mathematics and Statistics, University of Guelph, Ontario, N1G 2W1, Canada

\section*{A R T I C L E  I N F O}

Article history:
Received 10 December 2009
Available online 27 April 2010

AMS subject classifications:
primary 62H20
62G30
secondary 28A20

Keywords:
Maximum correlation
FGM distribution
Baker’s bivariate distribution
Order statistics
Fréchet–Hoeffding upper and lower bounds
Chebyshev’s inequality for integrals

\section*{A B S T R A C T}

We investigate Baker’s bivariate distributions with fixed marginals which are based on order statistics, and find conditions under which the correlation converges to the maximum for Fréchet–Hoeffding upper bound as the sample size tends to infinity. The convergence rate of the correlation is also investigated for some specific cases.

© 2010 Elsevier Inc. All rights reserved.

\section*{1. Introduction}

There are various ways to construct a bivariate distribution with fixed marginals. One of the most popular is the FGM distribution, dating back to Eyraud [7], Farlie [8], Gumbel [11] and Morgenstern [19]. Unlike the bivariate normal, which accommodates the full range \([-1, 1]\) of correlation, the FGM is rather restricted. Schucany, Parr and Boyer [21] pointed out that for continuous marginals the correlation in FGM is restricted to \([-1/3, 1/3]\). An ‘iterated’ version of FGM was proposed by Johnson and Kotz [15]. An initial iteration succeeded in enlarging the range to \([-1/3, 0.434]\) [13]. Subsequent iterations will no doubt extend the range further. Unfortunately no results are known to date due to the difficulty of determining the natural parameter space. Several other extensions of the FGM [14, 16, 19] all met with limited success.

A much wider class than the FGM is the Sarmanov family. There are Sarmanov densities with a correlation approaching 1 [23] for some (but not for all) marginals. To attain the maximum correlation, a Sarmanov density needs to concentrate all its mass on only two quadrants: \((x, y) : (x - x_0)(y - y_0) \geq 0\) for some real numbers \(x_0\) and \(y_0\). Some might argue that such distributions are rarely encountered in real life and that its applicability is limited (see also [3, 2] for generalized Sarmanov distributions).

Baker [4] has a novel approach. Let \(F_{k,n}\) be the distribution function of the \(k\)th smallest order statistic \(X_{k,n}\) of a random sample of size \(n\) from the marginal \(F\). Denote the corresponding quantities for \(G\) by \(G_{k,n}\) and \(Y_{k,n}\), where the two random samples from \(F\) and \(G\) are independent. Baker’s joint distribution is simply a linear combination of products of \(F_{k,n}\) and \(G_{k,n}\).

\* Corresponding author.
E-mail address: gdlin@stat.sinica.edu.tw (G.D. Lin).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.04.005
For maximal correlation she proposes (see Remark 1)

\[
H_+^{(n)}(x, y) = \frac{1}{n} \sum_{k=1}^{n} F_{k,n}(x)G_{k,n}(y),
\]

(1)

by choosing randomly one of the pairs \((X_{k,n}, Y_{k,n})\), \(k = 1, 2, \ldots, n\), and for the minimum,

\[
H_-^{(n)}(x, y) = \frac{1}{n} \sum_{k=1}^{n} F_{k,n}(x)G_{n-k+1,n}(y),
\]

(2)

by choosing randomly one of the pairs \((X_{k,n}, Y_{n-k+1,n})\), \(k = 1, 2, \ldots, n\). Clearly, both (1) and (2) satisfy the requirement of having the marginals \(F\) and \(G\). Unlike the Sarmanov, (1) and (2) have the same support as \(F \times G\) (and so does FGM), and allow \(F\) and \(G\) to be discrete distributions. Moreover, it seems reasonable that (1) possesses high correlation \(\rho_n\) at a large \(n\). Take, for example, \(F(x) = G(x) = 1 - e^{-x}, x > 0\). Shubina and Lee ([23], Example 3.1) proved that the maximum correlation attainable by the Sarmanov (with these marginals) is \(\rho = 0.6476\). For Baker’s (1), however, we see that \(\rho_n = 1 - \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k}\), which increases monotonically to 1, and at \(n = 8\), \(\rho_8 = 1479/2240 = 0.66027\) already exceeds the Sarmanov’s (for detail see the proof of Theorem 3(ii) in Section 3).

The purpose of this note is to investigate (i) the conditions under which the correlation for (1) converges to the limit and (ii) the convergence rate for some specific cases. In the next section we first review the upper and lower bounds of the correlation for all bivariate distributions with fixed marginals.

Remark 1. As a generalization of (1) and (2) Baker also introduces

\[
H_r^{(n)}(x, y) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} r_{k\ell} F_{k,n}(x)G_{\ell,n}(y),
\]

(3)

where \((r_{k\ell}) = r \in \mathcal{R}\) which consists of all \(r\) satisfying: \(r_{k\ell} \geq 0\) and \(\sum_{k=1}^{n} r_{k\ell} = \sum_{\ell=1}^{n} r_{k\ell} = 1/n\), for all \(k, \ell = 1, 2, \ldots, n\); namely, \(nr\) is a doubly stochastic matrix. We ask the question: Does (3) contain members with a correlation higher than that of (1)? The covariance of \(X\) and \(Y\) in (3) is \(\text{Cov}(X, Y) = \text{urv}^\top - \mu \nu\), where \(\mu = E(X), \nu = E(Y), \text{u} = (\mu_{1,n}, \mu_{2,n}, \ldots, \mu_{n,n}), \text{v} = (\nu_{1,n}, \nu_{2,n}, \ldots, \nu_{n,n})\) and \(\mu_{k,n} = E(X_{k,n}), \nu_{k,n} = E(Y_{k,n})\). We now claim that

\[
\max_{r \in \mathcal{R}} \text{urv}^\top \leq \text{ur}_0 \text{v}^\top
\]

where \(r_0 = \frac{1}{n} I_n\) and \(I_n\) is the \(n \times n\) identity matrix. In other words, the answer to the question is negative. To see this, note that \(\text{u}(nr)\) is majorized by \(\text{u} = \text{u}(nr_0)\), which in turn implies \(\text{urv}^\top \leq \text{ur}_0 \text{v}^\top\) for all \(r \in \mathcal{R}\) because of \(\nu_{1,n} \leq \nu_{2,n} \leq \cdots \leq \nu_{n,n}\) (see [17], p. 445).

2. The upper and lower bounds of correlations: the general case

Let \(H\) be any joint distribution of \(X\) and \(Y\) with the marginals \(F\) and \(G\). It is known that

\[
H_+(x, y) \leq H(x, y) \leq H_-(x, y) \quad \text{for } x, y \in R \equiv (-\infty, \infty),
\]

where \(H_+(x, y) \equiv \min[F(x), G(y)]\) and \(H_-(x, y) \equiv \max[F(x) + G(y) - 1, 0]\) are the Fréchet–Hoeffding upper and lower bounds, respectively [10,12]. This implies

\[
\text{Cov}_{H_-}(X, Y) \leq \text{Cov}_{H_0}(X, Y) \leq \text{Cov}_{H_+}(X, Y)
\]

by virtue of the Hoeffding representation \(\text{Cov}_{H}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [H(x, y) - F(x)G(y)] dx dy\). Equivalently,

\[
\rho_{H_-} \leq \rho_{H} \leq \rho_{H_+}.
\]

Let \(\sigma_X^2 = \text{Var}(X), \sigma_Y^2 = \text{Var}(Y)\) and let \(F^{-1}\) be the quantile function of \(F\), namely, \(F^{-1}(t) = \inf \{x : F(x) \geq t\}, t \in (0, 1)\). For convenience write \(F^{-1}(0) = \lim_{t \to 0^+} F^{-1}(t), F^{-1}(1) = \lim_{t \to 1^-} F^{-1}(t)\). Hereafter we shall consider only those cases where \(\sigma_X^2, \sigma_Y^2 \in (0, \infty)\). Some of the following properties about \(H_+\) and \(H_-\) have been mentioned in the works by Cambanis et al. [5] and Shih and Huang [22]. They are perhaps not as widely known as they should be, so we give a detailed proof here for the sake of completeness.

Proposition. Let \(X\) and \(Y\) be distributed by \(F\) and \(G\), respectively and let \(U\) be a uniform random variable independent of \(X\) and \(Y\). Then

(i) the joint distribution of \((F^{-1}(U), G^{-1}(U))\) is \(H_+\);
(ii) the joint distribution of \((F^{-1}(U), G^{-1}(1-U))\) is \(H_-\).


(iii) \( \rho_{H_+} = \rho^* \equiv (\int_0^1 F^{-1}(t)G^{-1}(t)dt - \mu \nu)/\sigma_X \sigma_Y \leq 1 \), and \( \rho^* = 1 \) if and only if the distributions \( F \) and \( G \) are of the same type, namely, \( F(x) = G(ax + b) \) on \( R \) for some \( a > 0 \) and \( b \in R \);
(iv) \( \rho_{H_-} = \rho \equiv (\int_0^1 F^{-1}(t)G^{-1}(1-t)dt - \mu \nu)/\sigma_X \sigma_Y \geq -1 \), and \( \rho_* = -1 \) if and only if the distributions of \( X \) and \( -Y \) are of the same type;
(v) \( \rho_* = -\rho^* \) if either \( F \) or \( G \) is a symmetric distribution.

**Proof.** (a) Note first that \( F^{-1}(t) \leq x \) if and only if \( t \leq F(x) \). We have thus, for \( x, y \in R \),

\[
P\{F^{-1}(U) \leq x, \ G^{-1}(U) \leq y\} = P\{U \leq F(x), \ U \leq G(y)\} = P\{U \leq \min\{F(x), \ G(y)\}\} = \min\{F(x), \ G(y)\} = H_+(x, y).
\]

(b) Similarly, we have, for \( x, y \in R \),

\[
P\{F^{-1}(U) \leq x, \ G^{-1}(1-U) \leq y\} = P\{U \leq F(x), \ 1-U \leq G(y)\} = P\{1-G(y) \leq U \leq F(x)\} = \max\{F(x) + G(y) - 1, \ 0\} = H_-(x, y).
\]

(c) To prove part (iii), note first that if the joint distribution of \( X \) and \( Y \) is \( H_+ \), then by (i) we have

\[
\text{Cov}_{H_+}(X, Y) = E(XY) - \mu \nu = \int_0^1 F^{-1}(t)G^{-1}(t)dt - \mu \nu
\]

and hence \( \rho_{H_+} = \rho^* \). Next, by the Cauchy inequality,

\[
\rho^* = \frac{1}{\sigma_X \sigma_Y} \int_0^1 (F^{-1}(t) - \mu)(G^{-1}(t) - \nu)dt
\]

\[
\leq \frac{1}{\sigma_X \sigma_Y} \left( \int_0^1 (F^{-1}(t) - \mu)^2dt \right)^{1/2} \left( \int_0^1 (G^{-1}(t) - \nu)^2dt \right)^{1/2} = 1,
\]

with the inequality reducing to equality if and only if \( F^{-1}(t) - \mu = a(G^{-1}(t) - \nu) \) on \( (0, 1) \) for some \( a > 0 \), namely, \( F \) and \( G \) are of the same type.

(d) Similarly, we have that \( \rho_{H_-} = \rho_* \) due to part (ii) and that

\[
-\rho_* = \frac{1}{\sigma_X \sigma_Y} \int_0^1 (F^{-1}(t) - \mu)(-G^{-1}(1-t) + \nu)dt
\]

\[
\leq \frac{1}{\sigma_X \sigma_Y} \left( \int_0^1 (F^{-1}(t) - \mu)^2dt \right)^{1/2} \left( \int_0^1 (-G^{-1}(1-t) + \nu)^2dt \right)^{1/2} = 1,
\]

with the inequality reducing to equality if and only if \( F^{-1}(t) - \mu = c(-G^{-1}(1-t) + \nu) \) on \( (0, 1) \) for some \( c > 0 \), namely, the distributions of \( X \) and \( -Y \) are of the same type because the random variable \( G^{-1}(1-U) \) is distributed as \( Y \).

(e) Without loss of generality, suppose that \( G \) is symmetric about 0. Then \( \nu = 0 \) and \( G^{-1}(t) = -G^{-1}(1-t) \) almost everywhere on \( (0, 1) \). This implies \( \rho_* = -\rho^* \). □

For the special case of \( F^{-1}(0) = G^{-1}(0) = 0 \), we have a more direct proof of \( \rho_{H_+} = \rho^* \). Write

\[
\text{Cov}_{H}(X, Y) = \int_0^\infty \int_0^\infty [P(X \leq x, \ Y \leq y) - F(x)G(y)]dxdy
\]

\[
= \int_0^\infty \int_0^\infty [P(X > x, \ Y > y) - (1-F(x))(1-G(y))]dxdy
\]

\[
= \int_0^\infty \int_0^\infty P(X > x, \ Y > y)dxdy - \mu \nu.
\]

Then, for \( H = H_+ \), we calculate

\[
\int_0^\infty \int_0^\infty P(X > x, \ Y > y)dxdy = \int_0^\infty \int_0^\infty [1-F(x) - G(y) + \min\{F(x), \ G(y)\}]dxdy
\]

\[
= \int_0^1 \int_0^1 (1-u - v + \min\{u, \ v\})dF^{-1}(u)dG^{-1}(v)
\]

\[
= \int_0^1 \left[ \int_0^u (1-v)dF^{-1}(u) + \int_v^1 (1-u)dF^{-1}(u) \right] dG^{-1}(v)
\]
This implies that \( H_+ \) has the correlation \( \rho^+ \). Similarly, we can prove that \( H_- \) has the correlation \( \rho_- \).

3. Convergence of \( \rho_n \) in Baker’s model

In this section we focus on Baker’s bivariate distributions. The product moment \( E(XY) \) for (1) is

\[
S_n = \frac{1}{n} \sum_{k=1}^{n} \left( \int_{-\infty}^{\infty} x f_{k,n}(x) \right) \left( \int_{-\infty}^{\infty} y g_{n}(y) \right) = \frac{1}{n} \sum_{k=1}^{n} \mu_{k,n} v_{k,n},
\]

while for (2) it becomes

\[
T_n = \frac{1}{n} \sum_{k=1}^{n} \mu_{k,n} v_{n-k+1,n}.
\]

Clearly, \( T_n \leq S_n \) for all \( n \) by Remark 1. Our first result of course follows immediately from the Proposition, but we give an alternative proof which is of interest in itself.

**Theorem 1.** (i) \( S_n \leq \int_{0}^{1} F^{-1}(t) G^{-1}(t) dt \) for each \( n \geq 1 \).

(ii) \( T_n \geq \int_{0}^{1} F^{-1}(t) G^{-1}(1-t) dt \) for each \( n \geq 1 \).

For the proof of the Theorem we use the following lemma whose proof can be found in [18], p. 246.

**Lemma 1** (Chebyshev’s Inequality for Integrals). Let \( f, g : (0, 1) \rightarrow R \) be Lebesgue integrable functions, both increasing or both decreasing. Furthermore, let \( p : (0, 1) \rightarrow (0, \infty) \) be an integrable function. Then

\[
\int_{0}^{1} p(x) f(x) dx \int_{0}^{1} p(x) g(x) dx \leq \int_{0}^{1} p(x) dx \int_{0}^{1} f(x) g(x) dx,
\]

provided that all integrals exist and are finite.

**Proof of Theorem 1.** (a) Note that \( E(X_{k,n}) = \int_{0}^{1} F^{-1}(t) f_{k,n}(t) dt \), where \( f_{k,n} \) is the Beta\( (k, n - k + 1) \) density function. Therefore,

\[
S_n = \frac{1}{n} \sum_{k=1}^{n} \left( \int_{0}^{1} F^{-1}(t) f_{k,n}(t) dt \right) \left( \int_{0}^{1} G^{-1}(t) f_{k,n}(t) dt \right)
\leq \frac{1}{n} \sum_{k=1}^{n} \left( \int_{0}^{1} F^{-1}(t) G^{-1}(t) f_{k,n}(t) dt \right) \left( \int_{0}^{1} f_{k,n}(t) dt \right)
= \int_{0}^{1} F^{-1}(t) G^{-1}(t) \frac{1}{n} \sum_{k=1}^{n} b_{k,n}(t) dt = \int_{0}^{1} F^{-1}(t) G^{-1}(t) dt,
\]

where the inequality is from **Lemma 1**. This proves part (i).

(b) To prove part (ii), note that \( b_{n-k+1,n}(t) = b_{k,n}(1-t) \) on \( (0, 1) \). Therefore,

\[
T_n = \frac{1}{n} \sum_{k=1}^{n} \left( \int_{0}^{1} F^{-1}(t) f_{k,n}(t) dt \right) \left( \int_{0}^{1} G^{-1}(1-t) f_{k,n}(t) dt \right)
\geq \frac{1}{n} \sum_{k=1}^{n} \left( \int_{0}^{1} F^{-1}(t) G^{-1}(1-t) f_{k,n}(t) dt \right) \left( \int_{0}^{1} f_{k,n}(t) dt \right)
= \int_{0}^{1} F^{-1}(t) G^{-1}(1-t) dt,
\]

where the inequality is from **Lemma 1** by letting \( p(t) = b_{k,n}(t) \), \( f(t) = F^{-1}(t) \) and \( g(t) = -G^{-1}(1-t) \). The proof is complete. \( \square \)
We next find conditions under which the correlation $\rho_n$ in (1) converges to the upper bound $\rho^*$, equivalently, $\lim_{n \to \infty} S_n = \int_0^1 F^{-1}(t)G^{-1}(t)dt$. The counterpart for the convergence of the correlation in (2) can be derived similarly. For the next theorem we need two more lemmas and the definition of two new classes of distributions.

**Lemma 2** (Pólya and Szegö [20], p. 52). Assume that the function $f$ is monotone and integrable on (0, 1). Then

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n-1} f \left( \frac{k}{n} \right) = \int_0^1 f(x)dx.
$$

**Lemma 3.** If $(X, Y)$ satisfies $\lim_{n \to \infty} S_n = \int_0^1 F^{-1}(t)G^{-1}(t)dt$ and if $H$ is the distribution of $Z = aX + b$ for some $a > 0$ and $b \in \mathbb{R}$, then $\lim_{n \to \infty} E(ZY) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} E(Z_{k,n})E(Y_{k,n}) = \int_0^1 H^{-1}(t)G^{-1}(t)dt$.

**Proof.** It is an immediate consequence of $EZ_{k,n} = aEX_{k,n} + b$ and $H^{-1}(t) = aF^{-1}(t) + b$, $t \in (0, 1)$. □

Let $\mathcal{L}$ and $\mathcal{U}$ be classes of distributions satisfying $E(X_{k,n}) \geq F^{-1}\left(\frac{k-1}{n}\right)$ all $(k, n)$ and $E(X_{k,n}) \leq F^{-1}\left(\frac{k}{n}\right)$ all $(k, n)$, respectively. These classes encompass many common distributions including the uniform, power, exponential, normal and some Gamma distributions. More generally, if $F^{-1}$ is convex on $(0, 1)$, then $F \in \mathcal{L}$ because, by Jensen’s inequality,

$$
E(X_{k,n}) = E(F^{-1}(U_{k,n})) \geq F^{-1}(E(U_{k,n})) = F^{-1}\left(\frac{k}{n+1}\right)
$$

where $U_{k,n}$ is the $k$th smallest order statistic of a random sample of size $n$ from the uniform distribution. Similarly, if $F^{-1}$ is concave on $(0, 1)$, then $F \in \mathcal{U}$. Moreover, if (a) the function $1/F$ is convex, or (b) $F$ is concave, or (c) $1/[1 - F]$ is concave, then $E(X_{k,n}) \geq F^{-1}\left(\frac{k-1}{n}\right)$ all $(k, n)$ and hence $F \in \mathcal{L}$. On the other hand, if (a) the function $F$ is convex, or (b) $1/F$ is concave, or (c) $1/[1 - F]$ is concave, or (d) $F$ is IFR (if $F$ has an increasing failure rate), then $F(E(X_{k,n})) \leq \frac{1}{n}$ for all $(k, n)$ and hence $F \in \mathcal{U}$ provided $F^{-1}$ is continuous on $(0, 1)$ (see, e.g., [6], Chapter 4).

**Theorem 2.** If either (i) $X \geq b, Y \geq c$ a.s. (almost surely) for some $b, c \in \mathbb{R}$ and $F, G \in \mathcal{L}$, or (ii) $X \leq b, Y \leq c$ a.s. for some $b, c \in \mathbb{R}$ and $F, G \in \mathcal{U}$, then $\lim_{n \to \infty} S_n = \int_0^1 F^{-1}(t)G^{-1}(t)dt$ and $\lim_{n \to \infty} \rho_n = \rho^*$.

**Proof.** By Lemma 3, it suffices to prove the case of $b = c = 0$ since both $\mathcal{L}$ and $\mathcal{U}$ are location-invariant.

(a) For part (i) with $b = c = 0$, the product moment

$$
S_n \geq \frac{1}{n} \sum_{k=1}^{n-1} F^{-1}\left(\frac{k-1}{n}\right)G^{-1}\left(\frac{k-1}{n}\right)
$$

$$
= \frac{1}{n} F^{-1}(0)G^{-1}(0) + \frac{1}{n} \sum_{k=2}^{n-1} F^{-1}\left(\frac{k-1}{n}\right)G^{-1}\left(\frac{k-1}{n}\right)
$$

$$
= \frac{1}{n} F^{-1}(0)G^{-1}(0) + \frac{1}{n} \sum_{k=1}^{n-1} F^{-1}\left(\frac{k}{n}\right)G^{-1}\left(\frac{k}{n}\right)
$$

$$
\to \int_0^1 F^{-1}(t)G^{-1}(t)dt \quad \text{as } n \to \infty,
$$

where the convergence is a consequence of the fact that the product $F^{-1}G^{-1}$ remains monotone and so Lemma 2 applies. Therefore, $\liminf_{n \to \infty} S_n \geq \frac{1}{n} F^{-1}(t)G^{-1}(t)dt$. This together with Theorem 1(i) implies that $\lim_{n \to \infty} S_n = \int_0^1 F^{-1}(t)G^{-1}(t)dt$ and hence $\lim_{n \to \infty} \rho_n = \rho^*$. This proves part (i).

(b) The proof for part (ii) is similar and is omitted. □

**Remark 2.** It is clear from the proof above that the condition in Theorem 2 can be weakened somewhat. For example, in proof (a) all we need is that for all large $n$, the average of $\mu_{k,n}v_{k,n}$ is greater than or equal to that of $F^{-1}\left(\frac{k-1}{n}\right)G^{-1}\left(\frac{k-1}{n}\right)$, instead of term-wise.

The following two examples show that the correlation $\rho_n$ for Baker’s (1) does converge to the upper bound $\rho^*$, which is less than 1.
Table 1

The correlation \( \rho_{n,p} \) for \( H_n^{(1)} \) with \( F(x) = x^p \) and \( G(y) = y^{1/2} \).

<table>
<thead>
<tr>
<th>n</th>
<th>( p )</th>
<th>1/20</th>
<th>1/5</th>
<th>1/2</th>
<th>1</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td>.363807</td>
<td>.567566</td>
<td>.642857</td>
<td>.645497</td>
<td>.579832</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>.472673</td>
<td>.722357</td>
<td>.801136</td>
<td>.792201</td>
<td>.696970</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>.605020</td>
<td>.903577</td>
<td>.977796</td>
<td>.949073</td>
<td>.813314</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>.618962</td>
<td>.922273</td>
<td>.995512</td>
<td>.964381</td>
<td>.824093</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td>.621623</td>
<td>.925834</td>
<td>.998876</td>
<td>.967278</td>
<td>.826121</td>
</tr>
<tr>
<td>( \infty )</td>
<td></td>
<td>.622514</td>
<td>.927025</td>
<td>1</td>
<td>.968246</td>
<td>.826797</td>
</tr>
</tbody>
</table>

Example 1 (Uniform & Exponential). Set \( F(x) = x, \ 0 < x < 1 \) and \( G(y) = 1 - e^{-y}, \ y > 0 \) in (1). Then we have \( \mu = \frac{1}{12}, \ \sigma_X^2 = \frac{1}{12}, \ \mu_{k,n} = \frac{k}{n+1} \) and \( \nu = 1, \ \sigma_Y^2 = 1, \ \nu_{k,n} = \sum_{j=n-k+1}^{n} \frac{1}{j} \). Therefore

\[
S_n = \frac{1}{n(n+1)} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{j}{1} (2n - j + 1) = \frac{1}{2n(n+1)} \sum_{j=1}^{n} (2n - j + 1)
\]

\[
= \frac{3n+1}{4n+1}
\]

\[
\implies \frac{3}{4} = \int_{0}^{1} t [-\ln(1-t)] \text{dt} = \int_{0}^{1} F^{-1}(t)G^{-1}(t) \text{dt} \quad \text{as} \quad n \to \infty.
\]

This implies that \( \lim_{n \to \infty} \rho_n = \rho^* = \sqrt{3}/2 \) for (1). For (2), we have \( \lim_{n \to \infty} T_n = 1/4 \) and \( \lim_{n \to \infty} \rho_n = \rho_* = -\sqrt{3}/2 = -\rho^* \) due to the symmetry of \( F \).

Example 2 (Power Function Marginals). Let \( F \) and \( G \) be the power function distributions with parameters \( p \) and \( 1/2 \), respectively, namely, \( F(x) = x^p, \ 0 < x < 1, \ p > 0 \) and \( G(y) = y^{1/2}, \ 0 < y < 1 \). It follows that \( F^{-1}(t) = t^{1/p}, \ \mu = \frac{p}{p+1}, \ \sigma_X^2 = \frac{p}{(p+1)^2(p+2)} \) and \( \mu_{k,n} = \frac{n! I(k+1/p)}{(n-k)! I(n+1/p)} \). Thus in (1) we have

\[
S_n = E(XY) = \frac{1}{n} \sum_{k=1}^{n} \frac{n! I(k+1/p)}{(k-1)! I(n+1/p)} \frac{k(k+1)}{(n+1)(n+2)}
\]

\[
\text{Cov}(X,Y) = \frac{1}{n} \sum_{k=1}^{n} \frac{n! I(k+1/p)}{(k-1)! I(n+1/p)} \frac{k(k+1)}{(n+1)(n+2)} - \frac{p}{3(p+1)}.
\]

\[
\rho_{n,p} = \sqrt{\frac{45(p+1)^2(p+2)}{4p} \left( \frac{1}{n} \sum_{k=1}^{n} \frac{n! I(k+1/p)}{(k-1)! I(n+1/p)} \frac{k(k+1)}{(n+1)(n+2)} - \frac{p}{3(p+1)} \right)}
\]

\[
\rho_{\infty,p} = \sqrt{\frac{45(p+1)^2(p+2)}{4p} \left( \int_{0}^{1} F^{-1}(t)G^{-1}(t) \text{dt} - \frac{p}{3(p+1)} \right)} = \sqrt{\frac{5p(p+2)}{3p+1}}.
\]

Table 1 gives the numerical results of \( \rho_{n,p} \) and \( \rho_{\infty,p} \). Note that \( \rho_{\infty,p} = 1 \) if and only if \( p = 1/2 \), namely when the two marginals are identical.

Finally, we investigate the convergence rates of \( \rho_n \) and obtain the following results.

Theorem 3. (i) If \( F \) is uniform and if \( G \) is an exponential distribution, then the convergence rate of \( \rho_n \) is \( 1/n \) as \( n \to \infty \).

(ii) If \( F = G \) is exponential, then the convergence rate of \( \rho_n \) is \((\ln n)/n\) as \( n \to \infty \).

Proof. Note first that the convergence rate of \( \rho_n \) is the same as that of \( S_n \) as \( n \to \infty \). So it suffices to calculate the convergence rates of \( S_n \).

(a) To prove part (i), we have, by Example 1,

\[
S_n = \frac{3n+1}{4(n+1)} = \frac{3}{4} - \frac{1}{2n} + O(n^{-2}) \quad \text{as} \quad n \to \infty.
\]
So the convergence rate of $S_n$ is $1/n$ as $n \to \infty$.

(b) To prove part (ii), we have $E(X_{k,n}) = \sum_{j=n-k+1}^{n} 1/j$ and hence

$$S_n = \frac{1}{n} \sum_{k=1}^{n} (EX_{k,n})^2 = \frac{1}{n} \sum_{k=1}^{n} \left( \sum_{j=n-k+1}^{n} \frac{1}{j} \right)^2$$

$$= \frac{1}{n} \sum_{k=1}^{n} \frac{2k - 1}{k} = 2 - \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k}$$

$$= 2 - \frac{\ln n}{n} + O(n^{-1}) \quad \text{as} \quad n \to \infty.$$ 

Therefore the convergence rate of $S_n$ is $(\ln n)/n$ as $n \to \infty$. □

Acknowledgments

The authors would like to thank the Editor and Referees for helpful comments. An original motivation of our paper is to point out that Baker’s claim of $\rho_n \to 1$ for her distribution (1) is false unless the two marginals are of the same type. After the submission of the paper, Professor Rose Baker was kind enough to call our attention to the Fréchet–Hoeffding bound and that no joint distribution, let alone hers, has $\rho_n \to 1$ if the marginals are not of the same type. Her remark led to an overhaul of our paper.

References