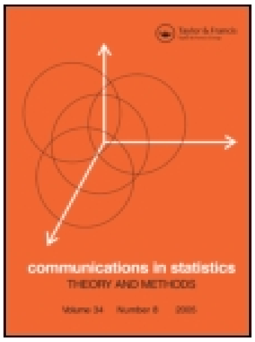

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Baker- Lin-Huang Type Bivariate Distributions Based on Order Statistics

K. Bayramoglu & I. Bayramoglu (Bairamov)

To cite this article: K. Bayramoglu & I. Bayramoglu (Bairamov) (2014) Baker- Lin-Huang Type Bivariate Distributions Based on Order Statistics, Communications in Statistics - Theory and Methods, 43:10-12, 1992-2006, DOI: 10.1080/03610926.2013.775301

To link to this article: https://doi.org/10.1080/03610926.2013.775301

Published online: 02 Jun 2014.

Communications in Statistics—Theory and Methods, 43: 1992–2006, 2014

ISSN: 0361-0926 print / 1532-415X online
DOI: 10.1080/03610926.2013.775301

Distribution Theory

Baker- Lin-Huang Type Bivariate Distributions Based on Order Statistics

K. BAYRAMOGLU[1] AND I. BAYRAMOGLU (BAIRAMOV)[2]

[1]Department of Statistics, Middle East Technical University, Ankara, Turkey
[2]Department of Mathematics, Izmir University of Economics, Izmir, Turkey

Baker (2008) introduced a new class of bivariate distributions based on distributions of order statistics from two independent samples of size n. *Lin and Huang (2010) discovered an important property of Baker's distribution and showed that the Pearson's correlation coefficient for this distribution converges to maximum attainable value, i.e., the correlation coefficient of the Fréchet upper bound, as* n *increases to infinity. Bairamov and Bayramoglu (2013) investigated a new class of bivariate distributions constructed by using Baker's model and distributions of order statistics from dependent random variables, allowing higher correlation than that of Baker's distribution. In this article, a new class of Baker's type bivariate distributions with high correlation are constructed based on distributions of order statistics by using an arbitrary continuous copula instead of the product copula.*

Keywords Bivariate distribution function; FGM distributions; Copula; Positive quadrant dependent; Negative quadrant dependent; Order statistics; Pearson's correlation coefficient

Mathematics Subject Classification 62H20; 62G30

1. Introduction

Huang and Kotz (1999) introduced new modifications of classical Farlie-Gumbel-Morgenstern (FGM) distribution introducing additional parameters. The new Huang-Kotz FGM distributions allow a correlation higher than the classical FGM and because of the simple analytical form aroused the interest of many researchers. In recent years, there appeared many articles dealing with the modifications of FGM distribution allowing high correlation. For related works on this subject, see Lai and Xie (2000), Bairamov et al. (2001), Amblard and Girard (2002), Bairamov and Kotz (2002, 2003), and Fisher and Klein (2007), among the others. Baker (2008) used a novel approach connected with the FGM distribution and introduces a new class of bivariate distributions based on the distributions of order statistics. Dolati and Úbeda-Flores (2009) considered new transformation of copulas based on Baker's construction, recovering known families of copulas. Baker

Received June 17, 2012; Accepted February 4, 2013.
Address correspondence to K. Bayramoglu, Department of Statistics, Middle East Technical University, Ankara, Turkey. E-mail: konul@metu.edu.tr

(2008) considered independent random variables X_i, Y_i , from two univariate distributions with distribution functions (cdf) F_X and F_Y, respectively. The corresponding probability density functions (pdf) are f_X and f_Y. Let $U_1 = \min(X_1, X_2)$ and $U_2 = \max(X_1, X_2)$, $V_1 = \min(Y_1, Y_2)$ and $V_2 = \max(Y_1, Y_2)$. To obtain positive correlation, Baker randomly chooses either the pair U_1, V_1 or U_2, V_2. The random numbers are now positively correlated, but the marginal distributions remain unchanged, because the random choice of either of two order statistics from a distribution gives a random variable from that distribution. To obtain a negative correlation, either U_1,V_2 or U_2, V_1 are chosen. The bivariate distribution of a randomly chosen pair of order statistics is

$$1/2 \times \left\{F_X^{2:2}(x)F_Y^{2:2}(y) + F_X^{1:2}(x)F_Y^{1.2}(y)\right\}$$

and by choosing either a pair of order statistics with probability q or the original independent random variables X, Y with probability $1 - q$, the bivariate distribution function is

$$\begin{aligned} H(x, y) &= (1 - q)F_X(x)F_Y(y) + (q/2)\left\{F_X^{2:2}(x)F_Y^{2:2}(y) + F_X^{1:2}(x)F_Y^{1:2}(y)\right\} \\ &= F_X(x)F_Y(y)\{1 + q(1 - F_X(x))(1 - F_Y(y)\}. \end{aligned}$$

In general, let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots Y_n$ are independent and identically distributed (i.i.d.) random variables with distribution functions (df) F_X and F_Y, respectively. Let $X_{k:n}$ and $Y_{k:n}$, $k = 1, 2, \ldots, n$ be corresponding order statistics and $F_X^{k:n}(x) = P\{X_{k:n} \le x\}$, $F_Y^{k:n}(y) = P\{Y_{k:n} \le y\}$. Baker's bivariate distribution function is now defined as

$$H_+^{(n)}(x, y) = \frac{1}{n}\sum_{k=1}^{n} F_X^{k:n}(x)F_Y^{k:n}(y) \tag{1}$$

$$H_-^{(n)}(x, y) = \frac{1}{n}\sum_{k=1}^{n} F_X^{k:n}(x)F_Y^{n-k+1:n}(y). \tag{2}$$

For Baker's bivariate distribution $H_+^{(n)}(x, y)$ with exponential marginals $F_X(x) = F_Y(x) = 1-e^{-x}$, $x > 0$, the Pearson's correlation coefficient is $\rho_n = 1-\frac{1}{n}\sum_{k=1}^{n}\frac{1}{k}$, which increases monotonely to 1.

As a generalization of (1) and (2), Baker also introduced

$$H_r^{(n)}(x, y) = \sum_{k=1}^{n}\sum_{l=1}^{n} r_{kl} F_X^{k:n}(x)F_Y^{l:n}(y), \tag{3}$$

where $r_{kl} \ge 0$ and $\sum_{k=1}^{n} r_{kl} = \sum_{l=1}^{n} r_{kl} = \frac{1}{n}$, for all $k, l = 1, 2, \ldots, n$. Lin and Huang (2010) proved that (3) does not contain members with a correlation higher than that of (1). That is why the best bivariate distribution with higher positive correlation among the members (3) is (1). Similar consideration holds true for the negative correlation. Lin and Huang (2010) investigated the conditions under which the correlation for (1) converges to the limit. In particular, they showed that if either (i) $X \ge b, Y \ge c$ a.s. for some $b, c \in \mathbb{R}$ and $E(X_{k:n}) \ge F_X^{-1}(\frac{k-1}{n})$ and $E(Y_{k:n}) \ge F_Y^{-1}(\frac{k-1}{n})$ for all (k, n), or (ii) $X \le b, Y \le c$ a.s. for some $b, c \in \mathbb{R}$ and $E(X_{k:n}) \le F_X^{-1}(\frac{k}{n})$ and $E(Y_{k:n}) \le F_Y^{-1}(\frac{k}{n})$ for all (k, n) then $\lim_{n\to\infty}\rho_n = \rho^*$, where ρ^* is the correlation coefficient of the Fréchet-Hoeffding upper bound $H_+(x, y) = \min(F_X(x), F_Y(y))$ (see Fréchet, 1951). The results presented in Lin and Huang (2010) makes Baker's distribution attractive.

Recently, Bairamov and Bayramoglu (2013) observed that if one uses the dependent random variables (X, Y) with positive quadrant dependent (PQD) joint distribution function $F(x, y)$, in the Baker's model instead of independent random variables, then the correlation increases, while for negative quadrant dependent $F(x, y)$ it decreases. More precisely, let $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ be a bivariate sample with joint distribution function $F(x, y) = C(F_X(x), F_Y(y))$. Bairamov and Bayramoglu (2013) considered the following bivariate distribution functions constructed on the basis of the Baker's idea:

$$K_+^{(n)}(x, y) = \frac{1}{n}\sum_{r=1}^{n} P\{X_{r:n} \le x, Y_{r:n} \le y\}$$

$$K_-^{(n)}(x, y) = \frac{1}{n}\sum_{r=1}^{n} P\{X_{r:n} \le x, Y_{n-r+1:n} \le y\},$$

where $X_{i:n}$ and $Y_{j:n}$ are the ith and jth order statistics constructed on the basis of bivariate observations (X_i, Y_i), $(i = 1, 2, ..., n)$ with joint distribution function $F(x, y) = P\{X_i \le x, Y_i \le y\}$ and marginal distribution functions $F_X(x) = F(x, \infty), F_Y(y) = F(\infty, y)$ so that $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$, $Y_{1:n} \le Y_{2:n} \le \cdots \le Y_{n:n}$. The joint df of $X_{r:n}$ and $Y_{s:n}$ is given in David (1981) (see also Arnold et al., 1992) as

$$P\{X_{r:n} \le x, Y_{s:n} \le y\} = \sum_{i=r}^{n}\sum_{j=s}^{n}\sum_{k=a}^{b} c(n, k; i, j) p_{11}^{k} p_{12}^{i-k} p_{21}^{j-k} p_{22}^{n-i-j+k}, \tag{4}$$

where

$$c(n, k; i, j) = \frac{n!}{k!(i-k)!(j-k)!(n-i-j+k)!},$$
$$a = \max(0, i + j - n), b = \min(i, j) \tag{5}$$

and

$$\begin{aligned} p_{11} &= F(x, y) \\ p_{12} &= F_X(x) - F(x, y) \\ p_{21} &= F_Y(y) - F(x, y) \\ p_{22} &= 1 - F_X(x) - F_Y(y) + F(x, y). \end{aligned} \tag{6}$$

Then,

$$K_+^{(n)}(x, y) = \frac{1}{n}\sum_{r=1}^{n}\sum_{i=r}^{n}\sum_{j=r}^{n}\sum_{k=a}^{b} c(n, k; i, j) p_{11}^{k} p_{12}^{i-k} p_{21}^{j-k} p_{22}^{n-i-j+k} \tag{7}$$

$$K_-^{(n)}(x, y) = \frac{1}{n}\sum_{r=1}^{n}\sum_{i=r}^{n}\sum_{j=n-r+1}^{n}\sum_{k=a}^{b} c(n, k; i, j) p_{11}^{k} p_{12}^{i-k} p_{21}^{j-k} p_{22}^{n-i-j+k}. \tag{8}$$

It is clear that the marginal distributions of $K_+^{(n)}(x, y)$ and $K_-^{(n)}(x, y)$ are again $F_X(x)$ and $F_Y(y)$, i.e., $K_+^{(n)}(x, \infty) = F_X(x)$ and $K_-^{(n)}(\infty, y) = F_Y(y)$. It is shown that for a PQD joint distribution function $F(x, y)$ the positive correlation of $K_+^{(n)}(x, y)$ is higher than that of $H_+^{(n)}(x, y)$ and for NQD $F(x, y)$ and the negative correlation of $K_-^{(n)}(x, y)$ is smaller than that of $H_-^{(n)}(x, y)$.

In this article, we consider a new class of bivariate distribution functions using Baker's construction and considering any copula $C(u, v)$ instead of product copula $C(u, v) = \Pi(u, v) = uv$. It follows that for this new class of distributions if $C(u, v)$ is PQD, i.e., $C(u, v) \geq uv$, for all $(u, v) \in [0, 1]^2$, then the Pearson's correlation coefficient is higher than that of Baker's distribution. Similarly, if the copula is NQD, i.e., $C(u, v) \leq uv$, for all $(u, v) \in [0, 1]^2$ then then the Pearson's correlation coefficient is smaller than that of Baker's distribution. Due to important contributions in Lin and Huang (2010) and Huang et al. (2013) Baker's distribution became very attractive. In this article, we call all modifications constructed on the Baker's idea Baker-Lin-Huang Type distribution, the distribution (7), (8) Baker's Type I BB and the new distributions introduced in this paper Baker's Type II BB distributions.

2. New bivariate Baker's Type II BB Distributions Based on a Copula Approach

Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots Y_n$ be i.i.d. random variables with df's F_X and F_Y, respectively. Let $X_{k:n}$ and $Y_{k:n}$, $k = 1, 2, \ldots, n$ be corresponding order statistics and $F_X^{k:n}(x) = P\{X_{k:n} \leq x\}$, $F_Y^{k:n}(y) = P\{Y_{k:n} \leq y\}$.

Recall that a two-dimensional copula is a function $C(x, y)$ from $[0, 1]^2 = [0, 1] \times [0, 1]$ to $[0, 1]$ with the properties:

1. $C(x, 0) = 0 = C(0, y)$, $C(x, 1) = x$ and $C(1, y) = y$;
2. for every x_1, x_2, y_1, y_2 such that $0 \leq x_1 < x_2 \leq 1$ and $0 \leq y_1 < y_2 \leq 1$

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0.$$

According to Sklar's Theorem, if $F_{X,Y}(x, y)$ is a joint distribution function with continuous marginal distributions $F_X(x)$ and $F_Y(y)$, then there exists a unique copula C such that $F(x, y) = C(F_X(x), F_Y(y))$. Theory and applications of copulas are well documented in Nelsen (2005), and Balakrishnan and Lai (2009). Let $C(u, v)$ be any copula. Consider

$$G_+^{(n)}(x, y) = \frac{1}{n} \sum_{k=1}^{n} C\left(F_X^{k:n}(x), F_Y^{k:n}(y)\right) \tag{9}$$

$$G_-^{(n)}(x, y) = \frac{1}{n} \sum_{k=1}^{n} C\left(F_X^{k:n}(x), F_Y^{n-k+1:n}(y)\right). \tag{10}$$

It follows from the properties of a copula that the marginal distributions of $G_+^{(n)}(x, y)$ and $G_-^{(n)}(x, y)$ are F_X and F_Y, respectively. In fact,

$$G_+^{(n)}(x, \infty) = \frac{1}{n} \sum_{k=1}^{n} C\big(F_X^{k:n}(x), 1\big) = \frac{1}{n} \sum_{k=1}^{n} F_X^{k:n}(x) = F_X(x)$$

$$G_+^{(n)}(\infty, y) = \frac{1}{n} \sum_{k=1}^{n} C\big(1, F_Y^{k:n}(y)\big) = \frac{1}{n} \sum_{k=1}^{n} F_Y^{k:n}(y) = F_Y(y).$$

Similarly, $G_-^{(n)}(x,\infty) = F_X(x)$ and $G_-^{(n)}(\infty, y) = F_Y(y)$. From now on, we will denote by ρ_H the Pearson's correlation coefficient between any random variables X and Y with joint distribution function $H(x, y)$. It is clear that if $C(u, v) = \Pi(u, v) = uv$ then $G_+^{(n)}(x, y) = H_+^{(n)}(x, y)$ and $G_-^{(n)}(x, y) = H_-^{(n)}(x, y)$. Since $C(F_X^{k:n}(x), F_Y^{k:n}(y))$ is a bivariate cdf (with marginals $F_X^{k:n}(x)$ and $F_Y^{k:n}(y)$) then $G_+^{(n)}(x, y)$ is obviously a bivariate cdf as a convex combination of bivariate cdf's. The copula used in construction (9) and (10) will be called the "kernel" copula for Baker's Type II BB distribution.

Theorem 2.1. *If $C(u, v)$ is PQD then $\rho_{G_+^{(n)}} \geq \rho_{H_+^{(n)}}$ and if $C(u, v)$ is NQD then $\rho_{G_-^{(n)}} \leq \rho_{H_-^{(n)}}$.*

Proof. Since $C(u, v) \geq uv$ for all $(u, v) \in [0, 1]^2$, then $C(F_X^{k:n}(x), F_Y^{k:n}(y)) \geq F_X^{k:n}(x)F_Y^{k:n}(y)$ for all $(x, y) \in \mathbb{R}^2$. From the Hoeffding's formula (see Hoeffding, 1940) for correlation coefficient one has

$$\begin{aligned}
\rho_{G_+^{(n)}} &= \frac{1}{\sqrt{Var(x)Var(Y)}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[G_+^{(n)}(x, y) - F_X(x)F_Y(y)\right]dxdy \\
&= \frac{1}{\sqrt{Var(x)Var(Y)}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\frac{1}{n}\sum_{k=1}^{n} C\left(F_X^{k:n}(x), F_Y^{k:n}(y)\right)\right. \\
&\quad \left. - F_X(x)F_Y(y)\right]dxdy \qquad (11)
\end{aligned}$$

$$\begin{aligned}
\rho_{H_+^{(n)}} &= \frac{1}{\sqrt{Var(x)Var(Y)}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[H_+^{(n)}(x, y) - F_X(x)F_Y(y)\right]dxdy \\
&= \frac{1}{\sqrt{Var(x)Var(Y)}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\frac{1}{n}\sum_{k=1}^{n} F_X^{k:n}(x)F_Y^{k:n}(y)\right. \\
&\quad \left. - F_X(x)F_Y(y)\right]dxdy \qquad (12)
\end{aligned}$$

and $\rho_{G_+^{(n)}} \geq \rho_{H_+^{(n)}}$. Similarly, $\rho_{G_-^{(n)}} \leq \rho_{H_-^{(n)}}$. □

Example 2.1. Baker's Type II BB distributions with FGM "kernel" copula. Let

$$C(u, v) = uv(1 + \alpha(1 - u)(1 - v)), (u, v) \in [0, 1]^2, -1 \leq \alpha \leq 1. \qquad (13)$$

Consider

$$G_+^{(n)}(x, y) = \frac{1}{n}\sum_{k=1}^{n} F_X^{k:n}(x)F_Y^{k:n}(y)(1 + \qquad (14)$$

$$+\alpha\left(1 - F_X^{k:n}(x)\right)\left(1 - F_Y^{k:n}(y)\right),\ 0 \leq \alpha \leq 1.$$

$$G_-^{(n)}(x, y) = \frac{1}{n}\sum_{k=1}^{n} F_X^{k:n}(x)F_Y^{n-k+1:n}(y)(1 + \qquad (15)$$

$$+\alpha\left(1 - F_X^{k:n}(x)\right)\left(1 - F_Y^{n-k+1:n}(y)\right),\ -1 \leq \alpha \leq 0.$$

Table 1
Correlation coefficients $\rho_{G_+^{(n)}}$ and $\rho_{G_-^{(n)}}$ with
$F(x, y) = xy(1 + \alpha(1 - x)(1 - y)), \alpha = 1, F_X(x) = x, F_Y(y) = y, 0 \leq x, y \leq 1$ and
$\rho_{K_+^{(n)}}, \rho_{K_-^{(n)}}$ and $\rho_{H_+^{(n)}}, \rho_{H_-^{(n)}}$ with *Uniform*(0, 1) marginals

n	2	4	6	8	10	12	15	20
$\rho_{G_+^{(n)}}$	0.5467	0.7258	0.8039	0.8475	0.8753	0.8945	0.9144	0.9348
$\rho_{K_+^{(n)}}$	0.5133	0.6915	0.7761	0.8247	0.8561	0.8779	0.9006	0.9241
$\rho_{H_+^{(n)}}$	0.3333	0.6000	0.7143	0.7778	0.8182	0.8462	0.8750	0.9048
$\rho_{G_-^{(n)}}$	−0.5467	−0.7258	−0.8039	−0.8475	−0.8753	−0.8945	−0.9144	−0.9348
$\rho_{K_-^{(n)}}$	−0.5133	−0.6915	−0.7761	−0.8247	−0.8561	−0.8779	−0.9006	−0.9241
$\rho_{H_-^{(n)}}$	−0.3333	−0.6000	−0.7143	−0.7778	−0.8182	−0.8462	−0.8750	−0.9048

Let $F_X(x) = x, 0 \leq x \leq 1$ and $F_Y(y) = y, 0 \leq y \leq 1$. Since the FGM copula (13) is PQD for $\alpha \geq 0$ and is NQD for $\alpha \leq 0$, then $\rho_{G_+^{(n)}} \geq \rho_{H_+^{(n)}}$ for $\alpha \geq 0$, $\rho_{G_-^{(n)}} \leq \rho_{H_-^{(n)}}$, for $\alpha \leq 0$. It is clear that if $\alpha = 0$ then $G_+^{(n)}(x, y) = H_+^{(n)}(x, y)$ and $G_-^{(n)}(x, y) = H_-^{(n)}(x, y)$. In Table 1 we present some numerical values of $\rho_{G_+^{(n)}}, \rho_{K_+^{(n)}}, \rho_{H_+^{(n)}}, \rho_{G_-^{(n)}}, \rho_{K_-^{(n)}}$ and $\rho_{H_-^{(n)}}$for *Uniform*(0, 1) marginals. The numerical calculations are made in MATLAB which is one of the commonly accepted packages for coding mathematical models since it has built-in functions for probability distributions and allows probabilistic and mathematical operations.

It can be observed from Table 1 that $\rho_{G_-^{(n)}} \leq \rho_{K_-^{(n)}} \leq \rho_{H_-^{(n)}} \leq \rho_{H_+^{(n)}} \leq \rho_{K_+^{(n)}} \leq \rho_{G_+^{(n)}}$.

In Table 2, the values of the correlation coefficients for $\rho_{G_+^{(n)}}, \rho_{K_+^{(n)}}\rho_{H_+^{(n)}}$, $\rho_{G_-^{(n)}}, \rho_{K_-^{(n)}}$ and $\rho_{H_-^{(n)}}$ for *Uniform*(0, 1) and *Exponential*(1) marginal distributions.

Again, from Table 2 we have $\rho_{G_-^{(n)}} \leq \rho_{K_-^{(n)}} \leq \rho_{H_-^{(n)}} \leq \rho_{H_+^{(n)}} \leq \rho_{K_+^{(n)}} \leq \rho_{G_+^{(n)}}$.

3. Copula Representation of Baker's Type II BB Distribution

Denote by $\Pi(t, s) = ts$, $(t, s) \in [0, 1]^2$ a product copula. Let $(X_i, Y_i), i = 1, 2, ..n$ be a bivariate sample with joint distribution function $F(x, y) = C(F_X(x), F_Y(y))$. Consider the

Table 2
Correlation coefficients $\rho_{G_+^{(n)}}$ and $\rho_{G_-^{(n)}}$ with $F(x, y) = xy(1 + \alpha(1 - x)(1 - y)), \alpha =$
$1, F_X(x) = x, 0 \leq x \leq 1, F_Y(y) = 1 - \exp(-y), y \geq 0$ and $\rho_{K_+^{(n)}}, \rho_{K_-^{(n)}}$ and $\rho_{H_+^{(n)}}, \rho_{H_-^{(n)}}$
with the same marginals

n	2	4	6	8	10	12	15	20
$\rho_{G_+^{(n)}}$	0.4811	0.6343	0.7001	0.7367	0.7600	0.7762	0.7929	0.8102
$\rho_{K_+^{(n)}}$	0.4426	0.5951	0.6682	0.7105	0.7379	0.7572	0.7772	0.7980
$\rho_{H_+^{(n)}}$	0.2886	0.5196	0.6185	0.6735	0.7085	0.7327	0.7577	0.7835
$\rho_{G_-^{(n)}}$	−0.4811	−0.6343	−0.7001	−0.7367	−0.7600	−0.7762	−0.7929	−0.8102
$\rho_{K_-^{(n)}}$	−0.4426	−0.5951	−0.6682	−0.7105	−0.7379	−0.7572	−0.7772	−0.7980
$\rho_{H_-^{(n)}}$	−0.2886	−0.5196	−0.6185	−0.6735	−0.7085	−0.7327	−0.7577	−0.7835

joint distribution of order statistics $X_{r:n}$ and $Y_{s:n}$ given in (4):

$$F_{X_{r:n},Y_{s:n}}(x, y) = P\{X_{r:n} \leq x, Y_{s:n} \leq y\}$$
$$= \sum_{i=r}^{n}\sum_{j=s}^{n}\sum_{k=a}^{b} c(n, k; i, j) p_{11}^{k} p_{12}^{i-k} p_{21}^{j-k} p_{22}^{n-i-j+k}, \tag{16}$$

where $c(n, k; i, j)$, $p_{11}, p_{12}, p_{21}, p_{22}$ are given in (5) and (6). The copula of bivariate distribution $F_{X_{r:n},Y_{s:n}}(x, y)$ is of considerable interest. Denote this copula as $C_{r,s:n}(t, s)$, then

$$F_{X_{r:n},Y_{r:n}}(x, y) = C_{r,s:n}(F_{X_{r:n}}(x), F_{Y_{s:n}}(y)). \tag{17}$$

It is well known (David, 1981), that

$$F_{X_{r:n}}(x) = \sum_{i=r}^{n}\binom{n}{i} F_X^i(x)(1 - F_X(x))^{n-i}$$
$$= \frac{1}{B(r, n-r+1)}\int_0^{F_X(x)} u^{r-1}(1-u)^{n-r}du = I_{r,n-r+1}(F_X(x))$$

and

$$F_{Y_{s:n}}(y) = \sum_{i=s}^{n}\binom{n}{i} F_Y^i(y)(1 - F_Y(y))^{n-i}$$
$$= \frac{1}{B(s, n-s+1)}\int_0^{F_Y(y)} u^{s-1}(1-u)^{n-s}du = I_{s,n-s+1}(F_Y(y)),$$

where $I_{a,b}(p) = \frac{1}{B(a,b)}\int_0^p u^{a-1}(1-u)^{b-1}du$ is an incomplete Beta function. Denote by $I_{a,b}^{-1}(p)$ the inverse of $I_{a,b}(p)$. Let $F_{X_{r:n}}(x) = I_{r,n-r+1}(F_X(x)) = t$ and $F_{Y_{s:n}}(y) = I_{s,n-s+1}(F_Y(y)) = s$. Then

$$x = F_X^{-1}\left(I_{r,n-r+1}^{-1}(F_X(t))\right) \text{ and } y = F_Y^{-1}\left(I_{s,n-s+1}^{-1}(F_Y(s))\right). \tag{18}$$

From (17), and (18), one has

$$C_{r,s:n}(t, s) = F_{X_{r:n},Y_{r:n}}\left(F_X^{-1}\left(I_{r,n-r+1}^{-1}(t)\right), F_Y^{-1}\left(I_{s,n-s+1}^{-1}(s)\right)\right). \tag{19}$$

Therefore, a copula of joint distribution of order statistics $X_{r:n}$ and $Y_{s:n}$ is $C_{r,s:n}(t, s)$ given in (19). In a special case if $r = s = n$ one has

$$C_{n,n:n}(t, s) = F^n\left(F_X^{-1}\left(I_{n,1}^{-1}(t)\right), F_Y^{-1}\left(I_{n,1}^{-1}(s)\right)\right) \tag{20}$$
$$= F^n\left(F_X^{-1}(t^{1/n}), F_Y^{-1}(s^{1/n})\right) = C^n(t^{1/n}, s^{1/n}),$$

since $I_{n,1}(t) = t^n$ and $I_{n,1}^{-1}(v) = v^{1/n}$. The copula (20) is an extreme value copula, i.e., a copula of componentwise maxima $\mathbf{X}_{(n)} = (X_{n.n}, Y_{n:n})$ of a given sample $\mathbf{X} = (X_1, X_2, ..., X_n)^T$ and $\mathbf{Y} = (Y_1, Y_2, ..., Y_n)^T$ with common distribution function $F(x, y)$ of (X, Y). The extreme value copula presents interest in insurance and finance in modeling of extreme events. The extreme events are more disastrous than any previously observed events. A bivariate copula C_{ext} is called an extreme value copula, if there exists a copula C such that for all $(t, s) \in [0, 1]^2$

$$\lim_{n\to\infty} C^n(t^{1/n}, s^{1/n}) = C_{ext}(t, s).$$

It is known that the class of extreme value copulas coincides with the class of max-stable copulas, i.e., a copula C_{ext} is an extreme value copula if and only if for all $n \in \mathbb{N}$

$$C_{ext}(t, s) = C^n_{ext}(t^{1/n}, s^{1/n})$$

for all $(t, s) \in [0, 1]^2$ (see Gudendorf and Segers, 2010; Haug et al., 2011).

Analogously, one obtains the copula of joint distribution of order statistics $X_{1:n}$ and $Y_{1:n}$ as

$$\begin{aligned} C_{1,n:n}(t, s) &= t + s - 1 \\ &\quad + \left[(1-t)^{1/n} + (1-s)^{1/n} - 1 \right. \\ &\quad \left. + F\left(F_X^{-1}(1-(1-t)^{1/n}), F_Y(1-(1-s)^{1/n})\right)\right]^n \\ &= t + s - 1 + [(1-t)^{1/n} + (1-s)^{1/n} - 1 \\ &\quad + C(1-(1-t)^{1/n}, 1-(1-s)^{1/n})]^n \end{aligned} \tag{21}$$

by noting that the joint distribution function of $X_{1:n}$ and $Y_{1:n}$ is

$$\begin{aligned} F_{X_{1:n},Y_{1:n}}(x, y) &= P\{X_{1:n} \le x, Y_{1:n} \le y\} \\ &= (1-(1-F_X(x))^n) + (1-(1-F_Y(y))^n) - 1 + \bar{F}^n(x, y) \\ &= (1-(1-F_X(x))^n) + (1-(1-F_Y(y))^n) - 1 \\ &\quad + (1 - F_X(x) - F_Y(y) + F(x, y))^n \\ &= C_{1,n:n}\left(1-(1-F_X(x)))^n, 1-(1-F_Y(y))^n\right). \end{aligned} \tag{22}$$

3.1. The Baker's Type II BB Distributions with "Kernel" Copula of Bivariate FGM Extreme Order Statistics

If the underlying distribution is classical FGM, i.e.,

$$F(x, y) = F_X(x)F_Y(y)(1 + \alpha(1 - F_X(x))(1 - F_Y(y))), -1 \le \alpha \le 1 \tag{23}$$

one obtains from (20)

$$C_{n:n}(t, s) = ts\left[1 + \alpha(1 - t^{1/n})(1 - s^{1/n})\right]^n \tag{24}$$

and from (21)

$$\begin{aligned} C_{1:n}(t, s) &= t + s - 1 + [(1-t)^{1/n} + (1-s)^{1/n} - 1 \\ &\quad + (1-(1-t)^{1/n})(1-(1-s)^{1/n}) \\ &\quad \times \{1 + \alpha(1-t)^{1/n}(1-s)^{1/n}\}]^n. \end{aligned} \tag{25}$$

It follows that $\lim_{n\to\infty} C_{n:n}(t,s) = ts = \Pi(t,s)$ and $\lim_{n\to\infty} C_{1:n}(t,s) = \Pi(t,s)$, since $\lim_{n\to\infty}[1+\alpha(1-t^{1/n})(1-s^{1/n})]^n = 1$ and

$$\begin{aligned}
&\lim_{n\to\infty}[(1-t)^{1/n} + (1-s)^{1/n} - 1 \\
&\quad + (1-(1-t)^{1/n})(1-(1-s)^{1/n})\{1+\alpha(1-t)^{1/n}(1-s)^{1/n}\}]^n \\
&= (1-t)(1-s),
\end{aligned}$$

where $\Pi(t,s)$ is a product copula of independent random variables. This means that if the joint distribution of (X, Y) is FGM given in (23), then the extreme order statistics $X_{n:n}$ and $Y_{n:n}$ are asymptotically independent. So are the order statistics $X_{1:n}$ and $Y_{1:n}$. In spite of this fact, the Baker's type BB distribution constructed on the base of the copula (24) and (25) has correlation large enough.

Indeed, consider Baker's type BB distribution (9) with the "kernel" copula (24) and with the "kernel" copula (25):

$$\begin{aligned}
\check{G}_+^{(n)}(x,y) &= \frac{1}{n}\sum_{k=1}^{n} C_{n,n:n}\left(F_X^{k:n}(x), F_Y^{k:n}(y)\right) \\
&= \frac{1}{n}\sum_{k=1}^{n} F_X^{k:n}(x)F_Y^{k:n}(y)\left[1+\alpha\left(1-\left[F_X^{k:n}(x)\right]^{1/n}\right)\left(1-\left[F_Y^{k:n}(y)\right]^{1/n}\right)\right]^n \\
\hat{G}_+^{(n)}(x,y) &= \frac{1}{n}\sum_{k=1}^{n} C_{1,n:n}\left(F_X^{k:n}(x), F_Y^{k:n}(y)\right) \\
&= \frac{1}{n}\sum_{k=1}^{n}\Big\{F_X^{k:n}(x) + F_Y^{k:n}(y) - 1 + \Big[\left(1-F_X^{k:n}(x)\right)^{1/n} + \left(1-F_Y^{k:n}(y)\right)^{1/n} - 1 \\
&\quad + \left(1-\left(1-F_X^{k:n}(x)\right)^{1/n}\right)\left(1-\left(1-F_Y^{k:n}(y)\right)^{1/n}\right) \\
&\quad \times \left\{1+\alpha\left(1-F_X^{k:n}(x)\right)^{1/n}\left(1-F_Y^{k:n}(y)\right)^{1/n}\right\}\Big]^n\Big\}.
\end{aligned}$$

For example, the correlation coefficient of $\check{G}_+^{(n)}(x,y)$ for *Uniform*(0, 1) marginals is $\rho_{\hat{G}_+^{(n)}} = 0.4985$ for $n = 2$.

3.2. The Baker's Type BB Distributions with "Kernel" Copula of Bivariate Gumbel's Extreme Order Statistics

Let

$$C(u,v) = \frac{uv}{1+u-uv}. \tag{26}$$

Equation (26) is the copula of the Gumbel's bivariate logistic distribution function

$$H_{X,Y}(x,y) = (1+e^{-x}+e^{-y})^{-1},\ x \geq 0,\ y \geq 0$$

with standard logistic marginal distributions $H_X(x) = (1+e^{-x})^{-1}$ and $H_Y(y) = (1+e^{-y})^{-1}$, $x \geq 0$, $y \geq 0$. (Gumbel, 1961; see also Nelsen, 2005, p. 28). Using

$$F(x, y) = \frac{F_X(x)F_Y(y)}{F_X(x) + F_Y(y) - F_X(x)F_Y(y)},$$

one obtains from (20)

$$C_{n:n}(t, s) = \frac{ts}{(t^{1/n} + s^{1/n} - t^{1/n}s^{1/n})^n} \tag{27}$$

and from (21)

$$\begin{aligned} C_{1:n}(t, s) = t + s - 1 + \\ + \Bigg[(1-t)^{1/n} + (1-s)^{1/n} - 1 \\ + \frac{(1-(1-t)^{1/n})(1-(1-s)^{1/n})}{2-(1-t)^{1/n}-(1-s)^{1/n}-(1-(1-t)^{1/n})(1-(1-s)^{1/n})}\Bigg]^n. \end{aligned} \tag{28}$$

It is seen that

$$\lim_{n\to\infty} C_{n:n}(t, s) = \lim_{n\to\infty} \frac{ts}{(t^{1/n} + s^{1/n} - t^{1/n}s^{1/n})^n} = ts = \Pi(t, s).$$

4. Joint Distribution of Bivariate Order Statistics for Fréchet Upper Bound Copula

Let $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ be a bivariate sample with joint distribution function $F(x, y) = C(F_X(x), F_Y(y))$. In this section our aim is first, to investigate the joint distribution function of bivariate order statistics $(X_{r:n}, Y_{r:n})$ for a copula with maximal correlation, i.e., the Fréchet upper bound. We are interested then in distribution function

$$K_+^{(n)}(x, y) = \frac{1}{n}\sum_{r=1}^{n}\sum_{i=r}^{n}\sum_{j=r}^{n}\sum_{k=a}^{b} c(n, k; i, j)p_{11}^k p_{12}^{i-k} p_{21}^{j-k} p_{22}^{n-i-j+k},$$

in the case where the marginal distributions are uniform and $C(u, v) = \min(u, v)$. Recall that the coefficients $c(n, k; i, j)$ and p_{11}, p_{12}, p_{21}, p_{22} are given in (6) and (5). Second, we consider a distribution introduced in (9) with uniform marginals, i.e.,

$$\begin{aligned} G_+^{(n)}(x, y) &= \frac{1}{n}\sum_{k=1}^{n} C\left(F_X^{k:n}(x), F_Y^{k:n}(y)\right) \\ &= \frac{1}{n}\sum_{r=1}^{n} C(F_{U_{r:n}}(x), F_{V_{r:n}}(y)), \end{aligned}$$

where $C(t, s) = \min(t, s)$ and $F_{U_{r:n}}(x) = \sum_{i=r}^{n}\binom{n}{i}x^i(1-x)^{n-i}$ and $F_{V_{r:n}}(y) = \sum_{i=r}^{n}\binom{n}{i}y^i(1-y)^{n-i}$.

Assume that marginal distributions are *Uniform*(0, 1), i.e., $F_X(x) = x, 0 \leq x \leq 1$, $F_Y(y) = y,\ 0 \leq y \leq 1$. Denote by $(U_i, V_i), i = 1, 2, ..., n$ the random sample from the bivariate distribution $C(u,v)$, $0 \le u \le 1$, $0 \le v \le 1$ and $U_{1:n} \le U_{2:n} \le \cdots \le U_{n:n}$, $V_{1:n} \le V_{2:n} \le \cdots \le V_{n:n}$ be corresponding order statistics. Then from (4) for $r = s$ one has

$$\begin{aligned} P\{U_{r:n} \le u, V_{r:n} \le v\} \\ &= \sum_{i=r}^{n}\sum_{j=r}^{n}\sum_{k=a}^{b} c(n,k;i,j) C^k(u,v)(u - C(u,v))^{i-k} \\ &\quad \times (v - C(u,v))^{j-k}(\bar{C}(u,v))^{n-i-j+k},\ 0 \le u \le 1, 0 \le v \le 1. \end{aligned} \tag{29}$$

and

$$\begin{aligned} K_+^{(n)}(u,v) = \frac{1}{n}\sum_{r=1}^{n}\sum_{i=r}^{n}\sum_{j=r}^{n}\sum_{k=a}^{b} c(n,k;i,j) C(u,v)^k (u - C(u,v))^{i-k} \\ (v - C(u,v))^{j-k}(\bar{C}(u,v))^{n-i-j+k}, 0 \le u \le 1, 0 \le v \le 1. \end{aligned} \tag{30}$$

Lemma 4.1. *The joint distribution of $U_{r:n}$ and $V_{r:n}$ can be represented as*

$$\begin{aligned} P\{U_{r:n} \le u, V_{r:n} \le v\} \\ &= \sum_{i=r}^{n}\binom{n}{i} C^i(u,v)\left[v - C(u,v) + \bar{C}(u,v)\right]^{n-i} \\ &\quad + \sum_{j=r}^{n}\binom{n}{j} C^j(u,v)\left[u - C(u,v) + \bar{C}(u,v)\right]^{n-j} \\ &\quad - \sum_{i=r}^{n}\binom{n}{i} C^i(u,v)(\bar{C}(u,v))^{n-i} \\ &\quad + \sum_{i=r}^{n}\sum_{j=r}^{n}\sum_{k\ne i\ne j} c(n,k;i,j) C(u,v)^k (u - C(u,v))^{i-k} \\ &\quad \times (v - C(u,v))^{j-k}(\bar{C}(u,v))^{n-i-j+k}. \end{aligned} \tag{31}$$

Theorem 4.1. *If $C(u,v) = \min(u,v)$, then*

$$\begin{aligned} P\{U_{r:n} \le u, V_{r:n} \le v\} \\ &= \begin{cases} \sum_{i=r}^{n}\binom{n}{i} u^i\,[1-u]^{n-i} & \text{if} \quad u \le v \\ \sum_{i=r}^{n}\binom{n}{i} v^i\,[1-v]^{n-i} & \text{if} \quad u > v \end{cases} \\ &= \begin{cases} P\{U_{r:n} \le u\} & \text{if} \quad u \le v \\ P\{V_{r:n} \le v\} & \text{if} \quad u > v \end{cases} \end{aligned}$$

Consider (31). Let $C(u, v) = \min(u, v)$. Then it is clear that $(u - C(u, v))^{i-k}(v - C(u, v))^{j-k} = 0$ for those k satisfying $k \neq i$ and $k \neq j$, because $u - C(u, v) = u - u = 0$ if $u \leq v$ and $v - C(u, v) = 0$ if $u > v$. Therefore the last term of (31) vanishes and we have from (31)

$$\begin{aligned} & P\{U_{r:n} \leq u, V_{r:n} \leq v\} \\ & \quad = \begin{cases} \displaystyle\sum_{i=r}^{n} \binom{n}{i} u^i \, [1-u]^{n-i} & \textit{if} \quad u \leq v \\ \displaystyle\sum_{i=r}^{n} \binom{n}{i} v^i \, [1-v]^{n-i} & \textit{if} \quad u > v \end{cases} \\ & \quad = \begin{cases} P\{U_{r:n} \leq u\} & \textit{if} \quad u \leq v \\ P\{V_{r:n} \leq v\} & \textit{if} \quad u > v \end{cases}. \end{aligned} \tag{32}$$

Consider now the Baker's Type I BB copula obtained from (32)

$$\begin{aligned} K_+^{(n)}(u, v) &= \frac{1}{n} \sum_{r=1}^{n} P\{U_{r:n} \leq u, V_{r:n} \leq v\} \\ &= \begin{cases} \displaystyle\frac{1}{n} \sum_{r=1}^{n} \sum_{i=r}^{n} \binom{n}{i} u^i \, [1-u]^{n-i} & \textit{if} \quad u \leq v \\ \displaystyle\frac{1}{n} \sum_{r=1}^{n} \sum_{i=r}^{n} \binom{n}{i} v^i \, [1-v]^{n-i} & \textit{if} \quad u > v \end{cases} \\ &= \begin{cases} u & \textit{if} \quad u \leq v \\ v & \textit{if} \quad u > v \end{cases} = W(u, v) = \min(u, v) \end{aligned}$$

is nothing but the Fréchet upper bound itself. In other words, the Baker's BB distribution with uniform marginals and underlying joint distribution which is the Fréchet upper bound generates the Fréchet upper bound.

Now consider the Baker's Type II BB distribution with the "kernel" copula, being Fréchet upper bound, i.e.,

$$G_+^{(n)}(x, y) = \frac{1}{n} \sum_{r=1}^{n} C(F_{U_{r:n}}(x), F_{V_{r:n}}(y)),$$

where $C(t, s) = \min(t, s)$ and $F_{U_{r:n}}(x) = \sum_{i=r}^{n} \binom{n}{i} x^i (1 - x)^{n-i}$ and $F_{V_{r:n}}(y) = \sum_{i=r}^{n} \binom{n}{i} y^i (1 - y)^{n-i}$. It follows that

$$\begin{aligned} G_+^{(n)}(x, y) &= \frac{1}{n} \sum_{r=1}^{n} \min(F_{U_{r:n}}(x), F_{V_{r:n}}(y)) \\ &= \frac{1}{n} \sum_{r=1}^{n} \min(F_{U_{r:n}}(x), F_{V_{r:n}}(y)) \end{aligned}$$

$$= \begin{cases} \dfrac{1}{n}\displaystyle\sum_{r=1}^{n}\sum_{i=r}^{n}\binom{n}{i}x^{i}(1-x)^{n-i} & \text{if} \quad x \le y \\ \dfrac{1}{n}\displaystyle\sum_{r=1}^{n}\sum_{i=r}^{n}\binom{n}{i}y^{i}(1-y)^{n-i}) & \text{if} \quad x > y \end{cases}$$
$$= \min(x, y).$$

Therefore, if in the constructions $K_{+}^{(n)}(x, y)$ and $G_{+}^{(n)}(x, y)$ with *Uniform*(0, 1) marginals, one uses the Fréchet upper bound copula $W(t, s)$, then obtained copula is again the Fréchet upper bound copula. Example 1 indicates that if in the same construction one uses FGM copula then different distributions with high correlation can be obtained. For different constructions satisfying conditions of the Example 2.1, we have the following set of inequalities for correlation coefficient:

$$\rho_M \le \rho_{G_{-}^{(n)}} \le \rho_{K_{-}^{(n)}} \le \rho_{H_{-}^{(n)}} \le \rho_{H_{+}^{(n)}} \le \rho_{K_{+}^{(n)}} \le \rho_{G_{+}^{(n)}} \le \rho_W,$$

where ρ_M is the correlation coefficient of the Fréchet lower bound $M(t, s) = \max(t + s - 1, 0)$.

Acknowledgements

The authors wish to thank the anonymous referee for interesting comments and suggestions which resulted in improvement of the presentation of this article.

References

Arnold, B., Balakrishnan, N., Nagaraja, H. N. (1992). *A First Course in Order Statistics*. New York: John Wiley & Sons.

Amblard, C., Girard, S. (2002). Symmetry and dependence properties within a semiparametric family of bivariate copulas. *Nonparametric. Statist.* 14(6):715–727.

Bairamov, I., Kotz, S. (2002). Dependence structure and symmetry of Huang-Kotz FGM distributions and their extensions. *Metrika* 56(1):55–72.

Bairamov, I., Kotz, S. (2003). On a new family of positive quadrant dependent bivariate distributions. *Int. Mathemat. J.* 3(11): 1247–1254.

Bairamov, I., Kotz, S., Bekçi, M. (2001). New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics. *J. Appl. Statist.* 28(5): 521–536.

Bairamov, I., Bayramoglu, K. (2013). From Huang-Kotz distribution to Baker's distribution. *J. Multivariate Anal.* 113: 106–115.

Baker, R. (2008). An order-statistics-based method for constructing multivariate distributions with fixed marginals. *J. Multivariate Anal.* 99: 2312–2327.

Balakrishnan, N., Lai, C. D. (2009). *Continuous Bivariate Distributions*. 2nd ed. New York: Springer.

David, H. (1981). *Order Statistics*. 2nd ed. New York: John Wiley.

Dolati, A., Úbeda-Flores, M. (2009). Constructing copulas by means of pairs of order statistics. *Kybernetika* 45: 992–1002.

Fischer, M., Klein, I. (2007). Constructing generalized FGM copulas by means of certain univariate distributions. *Metrika* 65: 243–260.

Fréchet, M. (1951). On correlation matrices with fixed margins (Sur les tableaux de corrélation dont les marges sont données.). *Ann. Univ. Lyon Sect. A* 14: 53–77.

Gudendorf, G., Segers, J. (2010). Extreme-value copulas. In: Jaworski, P., Durante, F., Härdle, W. K., Rychlik, T., Eds. *Proc. the workshop on Copula Theory and its Applications*. New York: Springer, pp. 127–146.

Haug, S., Klüppelberg, C., Peng, L. (2011). Statistical models and methods for dependence in insurance data. *J. Kor. Statist. Soc.* 40: 125–139.
Hoeffding, W. (1940). Scale-invariant correlation theory. (Maszstabinvariante Korrelationstheorie.). *Schr. Math. Inst. Univ. Berlin* 5: 181–233.
Huang, J. S., Kotz, S. (1999). Modifications of the Farlie Gumbel Morgenstern distributions. A tough hill to climb. *Metrika* 49: 135–145.
Huang, J. S., Dou, X., Kuriki, S., Lin, G. D. (2013). Dependence structure of bivariate order statistics from bivariate distributions with applications to Bayramoglu's distributions. *J. Multivariate Anal.* 114: 201–208.
Lai, C. D., Xie, M. (2000). A new family of positive quadrant dependent bivariate distributions. *Statist. Probab. Lett.* 46: 359–364.
Lin, G. D., Huang, J. S. (2010). A note on the maximum correlation for Baker's bivariate distributions with fixed marginals. *J. Multivariate Anal.* 101: 2227–2233.
Nelsen, R. (2005). *An Introduction to Copulas*. 2nd ed. New York: Springer.

Appendix

Proof of Lemma 4.1. Separating terms for summation in (29) for $k = i = j, k = i \neq j$ and $k = j \neq i$ we have

$$
\begin{aligned}
&P\{U_{r:n} \le u, V_{r:n} \le v\} \\
&\quad = \sum_{i=r}^{n} c(n,i,i,i) C^i(u,v)(\bar{C}(u,v))^{n-i} \\
&\qquad + \sum_{i=r}^{n} \sum_{j=i+1}^{n} c(n,i,i,j) C^i(u,v)(u - C(u,v))^{i-i}(v - C(u,v))^{j-i}(\bar{C}(u,v))^{n-i-j+i} \\
&\qquad + \sum_{j=r}^{n} \sum_{i=j+1}^{n} c(n,i,i,j) C^j(u,v)(u - C(u,v))^{i-j}(v - C(u,v))^{j-j}(\bar{C}(u,v))^{n-i-j+j} \\
&\qquad + \sum_{i=r}^{n} \sum_{j=r}^{n} \sum_{k \neq i \neq j} c(n,k;i,j) C(u,v)^k (u - C(u,v))^{i-k}(v - C(u,v))^{j-k} \\
&\qquad \times (\bar{C}(u,v))^{n-i-j+k} \\
&\quad = \sum_{i=r}^{n} \binom{n}{i} C^i(u,v)(\bar{C}(u,v))^{n-i} \\
&\qquad + \sum_{i=r}^{n} \sum_{j=i+1}^{n} \frac{n!}{i!(j-i)!(n-j)!} C^i(u,v)(v - C(u,v))^{j-i}(\bar{C}(u,v))^{n-j} \\
&\qquad + \sum_{j=r}^{n} \sum_{i=j+1}^{n} \frac{n!}{j!(i-j)!(n-i)!} C^j(u,v)(u - C(u,v))^{i-j}(\bar{C}(u,v))^{n-i} \\
&\qquad + \sum_{i=r}^{n} \sum_{j=r}^{n} \sum_{k \neq i \neq j} c(n,k;i,j) C(u,v)^k (u - C(u,v))^{i-k} \\
&\qquad \times (v - C(u,v))^{j-k}(\bar{C}(u,v))^{n-i-j+k}. \qquad (33)
\end{aligned}
$$

Consider the second summation in (33) and changing index in the inner sum as $j - i = k$ ($j = i + k,\ j = i + 1 \Longrightarrow k = 1,\ j = n \Longrightarrow k = n - i$) we have

$$\sum_{i=r}^{n} \sum_{j=i+1}^{n} \frac{n!}{i!(j-i)!(n-j)!} C^i(u,v)(v - C(u,v))^{j-i}(\bar{C}(u,v))^{n-j}$$

$$= \sum_{i=r}^{n} C^i(u,v) \frac{n!}{i!(n-i)!} \sum_{k=1}^{n-i} \frac{(n-i)!}{k!(n-i-k)!}(v - C(u,v))^k (\bar{C}(u,v))^{n-i-k}$$

$$= \sum_{i=r}^{n} \binom{n}{i} C^i(u,v) \left[\sum_{k=0}^{n-i} \frac{(n-i)!}{k!(n-i-k)!}(v - C(u,v))^k (\bar{C}(u,v))^{n-i-k} - (\bar{C}(u,v))^{n-i} \right]$$

$$= \sum_{i=r}^{n} \binom{n}{i} C^i(u,v) \left[v - C(u,v) + \bar{C}(u,v) \right]^{n-i} - \sum_{i=r}^{n} \binom{n}{i} C^i(u,v)(\bar{C}(u,v))^{n-i}. \quad (34)$$

Analogously, the third term in (33) can be written as

$$\sum_{j=r}^{n} \sum_{i=j+1}^{n} \frac{n!}{j!(i-j)!(n-i)!} C^j(u,v)(u - C(u,v))^{i-j}(\bar{C}(u,v))^{n-i}$$

$$= \sum_{j=r}^{n} \binom{n}{j} C^j(u,v) \left[u - C(u,v) + \bar{C}(u,v) \right]^{n-i} \quad (35)$$

$$- \sum_{i=r}^{n} \binom{n}{i} C^i(u,v)(\bar{C}(u,v))^{n-i}.$$

Taking into account (34) and (35) in (33) we have (31). □