Journal of Computational and Applied Mathematics 248 (2013) 1–14

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Some novel discrete distributions under fourfold sampling schemes and conditional bivariate order statistics

Ismihan Bayramoglu (Bairamov) [a,*], Gülder Kemalbay [b]

[a] Izmir University of Economics, Department of Mathematics, 35330, Balçova, Izmir, Turkey

[b] Yildiz Technical University, Department of Statistics, 34220, Esenler, Istanbul, Turkey

ARTICLE INFO

Article history:
Received 8 August 2012
Received in revised form 28 December 2012

Keywords:
Bivariate binomial distribution
Bivariate order statistics
Conditional distributions

ABSTRACT

This paper presents some novel trivariate discrete distributions that are obtained by modifying the bivariate binomial distribution. These distributions are important probability models for the development of conditional bivariate order statistics. The distributional properties of bivariate order statistics are studied and derived under the condition that certain values of the underlying random vectors (X, Y) are truncated and fall in the threshold set $\{(t, s) \in \mathbb{R}^2 : t \leq u, s \leq v\}, (u, v) \in \mathbb{R}^2$.

1. Introduction

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be independent copies of the bivariate random vector (X, Y) with joint distribution function $F_{X,Y}(x, y) = C(F_X(x), F_Y(y))$, where $C(u, v),\ (u, v) \in [0, 1]^2$ is the connecting copula. Denote by $X_{r:n}$ and $Y_{s:n}$ the rth and sth order statistics of $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$, respectively. The joint distribution of bivariate order statistics $(X_{r:n}, Y_{s:n})$ can be easily derived from the bivariate binomial distribution, which was first introduced by Aitken and Gonin [1] in connection with the fourfold sampling scheme. The bivariate binomial distribution can be described as follows: suppose that our population consists of two independent samples and each sample has two individuals, A, A^c and B, B^c, with probabilities $P(AB) = \pi_{11}, P(AB^c) = \pi_{12}, P(A^cB) = \pi_{21}$ and $P(A^cB^c) = \pi_{22}$, where $\sum_{ij} \pi_{ij} = 1$. Under random sampling with replacement n times, let ξ denotes the number of trials in which A appears and η denotes the number of trials in which B appears, respectively. The joint probability mass function of (ξ, η) is

$$P\{\xi = i, \eta = j\} = \sum_{k=\max(0,i+j-n)}^{\min(i,j)} \frac{n!}{k!(i-k)!(j-k)!(n-i-j+k)!} \pi_{11}^{k} \pi_{12}^{i-k} \pi_{21}^{j-k} \pi_{22}^{n-i-j+k}, \quad (1)$$

where $i = 0, 1, \ldots, n; j = 0, 1, \ldots, n$. Formula (1) can be easily explained: if in n trials, A appears together with B k times and together with B^c $i - k$ times, then B appears together with A^c $j - k$ times and B^c appears together with A^c $n - i - j + k$ times. The bivariate distribution given in (1) is called the bivariate binomial distribution. For some discussion of the bivariate and multivariate binomial distributions, see [2–10].

Recently, Bairamov and Gultekin [11] have considered the novel trivariate and quadrivariate distributions constructed on the basis of the bivariate binomial distribution. Note that the bivariate binomial distribution can be obtained from the multinomial distribution if one sets $AB = C_1, AB^c = C_2, A^cB = C_3, A^cB^c = C_4, P(C_1) = p_{11}, P(C_2) = p_{12}, P(C_3) = p_{21}$, and $P(C_4) = p_{22}$. If we denote by ζ_i the number of cases in which C_i occurs out of n repetitions, where $i = 1, 2, 3, 4$, then $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)$ is multinomial, $\xi_1 = \zeta_1 + \zeta_2$ and $\xi_2 = \zeta_1 + \zeta_3$.

* Corresponding author.
E-mail addresses: ismihan.bayramoglu@ieu.edu.tr (I. Bayramoglu), kemalbay@yildiz.edu.tr (G. Kemalbay).

0377-0427/$ – see front matter
doi:10.1016/j.cam.2013.01.013

Durante and Jaworski [12] considered the conditional distribution function of random variables (X, Y) given $(X, Y) \in \mathfrak{R}$, where $\mathfrak{R}$ is a Borel set in $\mathbb{R}^2$ with joint distribution function

$$H_{\mathfrak{R}}(x, y) = P\{X \le x, Y \le y \mid (X, Y) \in \mathfrak{R}\},$$

and using this conditional distribution, introduced a threshold copula. The threshold copula has interesting and important applications for studying the dependence among financial markets, especially regarding spatial contagion. For more recent result concerning threshold copulas and contagion, see [13–16]. For some interesting applications of order statistics and their concomitants, bivariate distributions and copulas, in insurance, see [17–19].

In this work, we consider the joint distribution of bivariate order statistics $(X_{r:n}, Y_{s:n})$ under the condition that h of the random observations $(X_1, Y_1), \dots, (X_n, Y_n)$ are truncated, i.e., they fall in the set $\mathbf{B}_{uv} = \{(t, s) \in \mathbb{R}^2 : t \le u, s \le v\}$, $(u, v) \in \mathbb{R}^2$, assuming $P\{(X, Y) \in \mathbf{B}_{uv}\} > 0$. This conditional distribution is derived using novel modifications of the bivariate binomial distribution introduced in Section 2 of this paper. The results obtained in this paper have applications for studying the dependence among financial markets in crises or other extreme situations. The conditional bivariate order statistics can also be used in reliability analyses for studying the mean residual life functions of complex systems.

The statistical theory of reliability considers systems that consist of n components, and the lifetimes of these components are assumed to be nonnegative random variables. Recently, Bairamov [20] considered complex systems that consist of n elements, which each contain two or more components, and studied the reliability properties of such systems. Let a system consists of n elements, and assume that each element has two components, (A_i, B_i), $i = 1, 2, \dots, n$. Let X_i be the lifetime of the component A_i and Y_i be the lifetime of the component B_i, $i = 1, 2, \dots, n$. Then, (X_i, Y_i) represents the lifetime of the ith element. Assume that the components of the ith element are dependent, i.e., X_i and Y_i are dependent random variables with joint distribution function $F(x, y)$. As an example, Bairamov [20] considered (r, s)-out-of-n systems, which function if and only if at least r of the n components $A_1, A_2, \dots, A_n$ and s of the n components $B_1, B_2, \dots, B_n$ function. Then, the reliability of such a system is

$$P\{T > t\} = P\{X_{n-r+1:n} > t, Y_{n-s+1:n} > t\},$$

where T is the lifetime of the system and $(X_{r:n}, Y_{s:n})$ is the vector of bivariate rth and sth order statistics constructed from the sample $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. The mean residual life function of an (r, s)-out-of-n system with intact components at time t is

$$\begin{aligned}\Phi_{r,s:n}(t) &= E\{T - t \mid X_{1:n} > t, Y_{1:n} > t\}\\ &= E\{T^{(t)}_{r,s:n}\},\end{aligned}$$

where $T^{(t)}_{r,s:n}$ is a conditional random variable defined as $T^{(t)}_{r,s:n} \equiv (X_{n-r+1:n} - t, Y_{n-s+1:n} - t \mid \{$ none of the components has failed at time $t\})$. It is clear that to evaluate $\Phi_{r,s:n}(t)$, we must know the survival function of the conditional random variable $T^{(t)}_{r,s:n}$, i.e., the survival function of conditional order statistics, which is the subject of the present paper.

This paper is organised as follows: In Section 2, we consider novel trivariate distributions obtained from bivariate binomial distributions by introducing new events in a fourfold model. In Section 3, using the modified trivariate distributions we introduce, the conditional distributions of bivariate order statistics $(X_{r:n}, Y_{s:n})$, $1 \le r, s \le n$ constructed from bivariate observations (X_i, Y_i), $i = 1, 2, \dots, n$ are derived, where we assume that a certain number of these observations are truncated, i.e., fall in the threshold set $\{(t, s) \in \mathbb{R}^2, t \le u, s \le v\}$, where $(u, v) \in \mathbb{R}^2$.

1.1. The modified binomial distribution

Consider a fourfold sampling scheme, i.e., suppose that the outcome of the random experiment is one of the events A or A^c and simultaneously one of B or B^c with the probabilities $P(AB) = \pi_{11}$, $P(AB^c) = \pi_{12}$, $P(A^cB) = \pi_{21}$ and $P(A^cB^c) = \pi_{22}$, where $\sum_{ij} \pi_{ij} = 1$. More precisely, in this scheme, the event A occurs together with B or B^c and the event B occurs together with A or A^c. Therefore, the possible outcomes of the experiment are AB, AB^c, A^cB and A^cB^c. We will refer to this sampling scheme as the fourfold sampling scheme. We may also refer to this sampling scheme as a fourfold experiment. If we repeat the fourfold experiment independently n times, then we will use the expression "in n independent fourfold trials" or "in n independent trials of the fourfold experiment".

In this fourfold experiment setup, for further modifications of the bivariate binomial distribution, we consider the following four cases:

1. Together with A, B, A^c, B^c, the event C can also occur in the experiment, where $C \subset AB$.
2. The events C and D can also occur, where $C \subset AB$ and $D \subset AB^c$.
3. We assume that the events C and E can also occur, where $C \subset AB$ and $E \subset A^cB$.
4. The events D, E and F can also occur, where $D \subset AB^c$, $E \subset A^cB$ and $F \subset A^cB^c$.

Note that these four cases describe different situations and must be considered separately.

According to these four cases, we consider n independent trials of the fourfold experiment and define the random variables ξ, η and ζ as follows:

Definition 1. (a) If $C \subset AB$, then ξ, η and ζ are the number of occurrences of the events A, B, C, respectively.
(b) Let $C \subset AB$ and $D \subset AB^c$. Denote by ξ, η, ζ the number of occurrences of the events $A, B, C \cup D$, respectively.
(c) In the case in which, $C \subset AB$ and $E \subset A^cB$, we denote by ξ, η and ζ the number of occurrences of the events $A, B, C \cup E$, respectively.
(d) For $D \subset AB^c$, $E \subset A^cB$ and $F \subset A^cB^c$, the random variables ξ, η and ζ denote the number of occurrences of the events $A, B, D \cup E \cup F$, respectively.
(d1) If $D \subset AB^c$, $E \subset A^cB$ and $F \subset A^cB^c$, then ξ, η and ζ are the number of occurrences of the events $A, B, AB \cup D \cup E \cup F$, respectively.

Note that, the events C, D, E and F are distinct for each case (a), (b), (c), (d) and (d1) and ξ, η, ζ denote distinct random variables for each case, i.e., ξ, η and ζ in (a) are distinct from ξ, η and ζ in the other cases. We prefer to use such notation to avoid introducing a tremendous number of letters. Therefore, each of the cases (a), (b), (c), (d) and (d1) must be considered separately. The joint distributions of the random variables ξ, η and ζ for each of the cases (a), (b), (c), (d) and (d1) are given in the following Theorems 1–4.1.

Theorem 1. *In the fourfold sampling scheme, let $C \subset AB$ and ξ, η, ζ be the number of occurrences of the events A, B, C in n independent trials, respectively (case* (a) *in Definition* 1*). Then, the joint probability mass function of ξ, η and ζ is*

$$\begin{aligned} P_1(i,j,h) &\equiv P\{\xi = i, \eta = j, \zeta = h\} \\ &= \sum_{k=a}^{b} C_1(n; h, k, i, j) P(C)^h [P(AB) - P(C)]^{k-h} P(AB^c)^{i-k} P(A^cB)^{j-k} P(A^cB^c)^{n-i-j+k}, \end{aligned} \tag{2}$$

where

$$\begin{aligned} &C_1(n; h, k, i, j) = \frac{n!}{h!(k-h)!(i-k)!(j-k)!(n-i-j+k)!}; \\ &a = \max(0, i+j-n); \qquad b = \min(i,j); \quad i,j = 0, 1, \ldots, n; \\ &h = 0, \ldots, \min(i,j). \end{aligned}$$

Proof. If $\xi = i$, we consider all possible cases of the occurrence of the event A and we indicate these cases as $k = 0, 1, \ldots,$ then A occurs together with B k times and together with B^c $i-k$ times. $\zeta = h$ indicates that C occurs h times. Because $C \subset AB$, h may be at most $\min(i,j)$ because $\xi = i, \eta = j$. Then, $AB \setminus C = AB \cap C^c$ occurs $k-h$ times. $\eta = j$ implies that if B appears together with A^c $j-k$ times, B^c appears together with A^c $n-i-j+k$ times. Schematically, this situation can be described as follows:

$A \setminus B$	B	B^c
A	C h times; AB k times	AB^c $i-k$ times
A^c	A^cB $j-k$ times	A^cB^c $n-i-j+k$ times

Therefore, it is clear that if we repeat the experiment n times, then h outcomes of the event C can be observed in $\binom{n}{h}$ ways and $k-h$ outcomes of the event $AB \setminus C$ can be realised in $\binom{n-h}{k-h}$ ways. Then, $i-k$ outcomes of the event A can be observed with B^c in $\binom{n-h-(k-h)}{i-k} = \binom{n-k}{i-k}$ ways and A^c can be realised together with B in $\binom{n-k-(i-k)}{j-k} = \binom{n-i}{j-k}$ ways.

Thus in n independent trials, the number of possible cases in which A appears i times, B appears j times and C appears h times is

$$\binom{n}{h}\binom{n-h}{k-h}\binom{n-k}{i-k}\binom{n-i}{j-k} = \frac{n!}{h!(k-h)!(i-k)!(j-k)!(n-i-j+k)!}$$

with probability,

$$P(C)^h [P(AB) - P(C)]^{k-h} P(AB^c)^{i-k} P(A^cB)^{j-k} P(A^cB^c)^{n-i-j+k}.$$

It is clear that $\max(0, i+j-n) \le k \le \min(i,j)$ and $i,j = 0, 1, \ldots, n$; $h = 0, \ldots, \min(i,j)$. □

Remark 1. If $C = AB$, then ξ, η, ζ are the number of occurrences of the events A, B, AB in n independent trials, respectively. In this case, from (2) we have

$$P\{\xi = i, \eta = j, \zeta = h\} = \frac{n!}{h!(i-h)!(j-h)!(n-i-j+h)!} P(AB)^h P(AB^c)^{i-h} P(A^cB)^{j-h} P(A^cB^c)^{n-i-j+h}, \tag{3}$$
$$i, j = 0, 1, \ldots, n; \qquad h = \max(0, i+j-n), \ldots, \min(i, j),$$

and (1) is the marginal probability mass function (p.m.f.) of (3).

Theorem 2. *Consider the fourfold sampling scheme and assume that $C \subset AB$ and $D \subset AB^c$. Let ξ, η and ζ be the number of occurrences of the events $A, B, C \cup D$ in n independent trials, respectively (case* (b) *in Definition 1). Then, the joint probability mass function of ξ, η and ζ is*

$$\begin{aligned} P_2(i, j, h) &\equiv P\{\xi = i, \eta = j, \zeta = h\} \\ &= \sum_{k=a}^{b} \sum_{l=0}^{h} C_2(n; k, l, h, i, j) P(C)^l [P(AB) - P(C)]^{k-l} P(D)^{h-l} \\ &\quad \times [P(AB^c) - P(D)]^{i-k-h+l} P(A^cB)^{j-k} P(A^cB^c)^{n-i-j+k}, \end{aligned} \tag{4}$$

where

$$C_2(n; k, l, h, i, j) = \frac{n!}{l!(k-l)!(h-l)!(i-k-h+l)!(j-k)!(n-i-j+k)!};$$
$$a = \max(0, i+j-n); \qquad b = \min(i, j); \quad i, j = 0, 1, \ldots, n;$$
$$h = 0, 1, \ldots, i.$$

Proof. We know the implications of $\xi = i$ and $\eta = j$ from the proof of Theorem 1. Unlike in the previous theorem, $\zeta = h$, i.e., $C \cup D$ occurs h times. Because $C \subset AB$ and $D \subset AB^c$, $C \cup D \subset AB \cup AB^c = A$. Therefore, h can be at most i because $\xi = i$. Then, indicating all possible cases of the occurrence of event C by $l = 0, 1, 2 \ldots$, one observes that D occurs $h - l$ times. Hence, $AB \setminus C$ occurs $k - l$ times and $AB^c \setminus D$ occurs $i - k - (h - l)$ times. Then, similar to the proof of Theorem 1, all possible cases of the occurrence of the event $\{\xi = i, \eta = j, \zeta = h\}$ can be schematically described as follows:

$A \setminus B$	B	B^c
A	[C: l times] AB: k times	[D: $h - l$ times] AB^c: $i - k$ times
A^c	A^cB: $j - k$ times	A^cB^c: $n - i - j + k$ times

Then, in n independent repeated trials, l outcomes of the event C can be observed in $\binom{n}{l}$ ways and $k - l$ outcomes of the event $AB \setminus C$ can be realised in $\binom{n-l}{k-l}$ ways. Therefore, $h - l$ outcomes of the event D can be realised in $\binom{n-l-(k-l)}{h-l} = \binom{n-k}{h-l}$ ways and $i - k - h + l$ outcomes of the event $AB^c \setminus D$ can be realised in $\binom{n-k-(h-l)}{i-k-h+l}$ ways. Then, A^c can be realised together with B in $\binom{n-k-h+l-(i-k-h+l)}{j-k} = \binom{n-i}{j-k}$ ways.

Thus in n independent trials, the number of possible cases in which A appears i times, B appears j times and $C \cup D$ appears h times is

$$\binom{n}{l}\binom{n-l}{k-l}\binom{n-k}{h-l}\binom{n-k-(h-l)}{i-k-h+l}\binom{n-i}{j-k} = \frac{n!}{l!(k-l)!(h-l)!(i-k-h+l)!(j-k)!(n-i-j+k)!}$$

and each case has the same probability,

$$P(C)^l [P(AB) - P(C)]^{k-l} P(D)^{h-l} [P(AB^c) - P(D)]^{i-k-h+l} P(A^cB)^{j-k} P(A^cB^c)^{n-i-j+k}.$$

It is clear that $\max(0, i+j-n) \le k \le \min(i, j)$ and $i, j = 0, 1, \ldots, n;\ h = 0, 1, \ldots, i$. □

Theorem 3. *Let $C \subset AB$ and $E \subset A^cB$ in the fourfold sampling scheme. Assume that ξ, η and ζ denote the number of occurrences of the events $A,\ B, C \cup E$ in n independent trials, respectively (case* (c) *in Definition 1). Then, the joint probability mass function* *of* ξ, η *and* ζ *is*

$$\begin{aligned} P_3(i,j,h) &\equiv P\{\xi = i, \eta = j, \zeta = h\} \\ &= \sum_{k=a}^{b}\sum_{l=0}^{h} C_3(n;k,l,h,i,j)P(C)^l[P(AB)-P(C)]^{k-l}P(AB^c)^{i-k} \\ &\quad \times P(E)^{h-l}[P(A^cB)-P(E)]^{j-k-h+l}P(A^cB^c)^{n-i-j+k}, \end{aligned} \tag{5}$$

where

$$C_3(n;k,l,h,i,j) = \frac{n!}{l!(k-l)!(i-k)!(h-l)!(j-k-h+l)!(n-i-j+k)!};$$
$a = \max(0, i+j-n)$; $b = \min(i,j)$; $i, j = 0, 1, \ldots, n$;
$h = 0, 1, \ldots, j$.

Proof. This theorem can be proved in a manner similar to the proof of Theorem 2 using the below schematic representation:

$A \setminus B$	B	B^c
A	[C: l times] AB: k times	AB^c: $i-k$ times
A^c	[E: $h-l$ times] A^cB: $j-k$ times	A^cB^c: $n-i-j+k$ times

□

Theorem 4. *In the fourfold sampling scheme, let* $D \subset AB^c$, $E \subset A^cB$, $F \subset A^cB^c$ *and* ξ, η, ζ *be the number of occurrences of the events* A, B, $D \cup E \cup F$ *in* n *independent trials, respectively (case* (d) *in Definition 1). Then, the joint probability mass function of* ξ, η *and* ζ *is*

$$\begin{aligned} P_4(i,j,h) &\equiv P\{\xi = i, \eta = j, \zeta = h\} \\ &= \sum_{k=a}^{b}\sum_{p=0}^{i-k}\sum_{q=0}^{j-k} C_4(n;k,p,q,h,i,j)P(AB)^k \times P(D)^p[P(AB^c)-P(D)]^{i-k-p}P(E)^q \\ &\quad \times [P(A^cB)-P(E)]^{j-k-q} \times P(F)^{h-p-q}[P(A^cB^c)-P(F)]^{n-i-j+k-h+p+q}, \end{aligned} \tag{6}$$

where

$$C_4(n;k,p,q,h,i,j) = \frac{n!}{k!p!(i-k-p)!q!(j-k-q)!(h-p-q)!} \times \frac{1}{(n-i-j+k-h+p+q)!};$$
$a = \max(0, i+j-n)$; $b = \min(i,j)$; $i, j, h = 0, 1, \ldots, n$.

Proof. The schematic representation for this theorem is as follows:

$A \setminus B$	B	B^c
A	k times	[D: p times] AB^c: $i-k$ times
A^c	[E: q times] A^cB: $j-k$ times	[F: $h-p-q$ times] A^cB^c: $n-i-j+k$ times

For clarity of explanation, we denote by $\mu(M)$ the number of occurrence of any event M in n independent trials of the fourfold experiment. Because $D \cup E \cup F$ occurs h times, i.e., $\zeta = h$ and $D \cap E \cap F = \varnothing$, $h = \mu(D) + \mu(E) + \mu(F)$, where

$\mu(D) = p, \mu(E) = q, \mu(F) = h - p - q$ are the number of occurrences of the events D, E and F, respectively. Then, the number of occurrences of AB is k, of $AB^c \setminus D$ is $i - k - p$, of $A^cB \setminus E$ is $j - k - q$ and of $A^cB^c \setminus F$ is $n - i - j + k - (h - p - q)$.

The implications of $\xi = i$ and $\eta = j$ are also known from the proof of the first theorem.

Therefore, it is clear that if we repeat the experiment n times, then k outcomes of the event AB can be observed in $\binom{n}{k}$ ways, p outcomes of the event D can be observed in $\binom{n-k}{p}$ ways and $i - k - p$ outcomes of the event $AB^c \setminus D$ can be realised in $\binom{n-k-p}{i-k-p}$ ways. Then, q outcomes of the event E can be observed in $\binom{n-k-p-(i-k-p)}{q} = \binom{n-i}{q}$ ways and $j - k - q$ outcomes of the event $A^cB \setminus E$ can be realised in $\binom{n-i-q}{j-q-k}$ ways. Finally, $h - p - q$ outcomes of the event F can be realised in $\binom{n-i-q-(j-k-q)}{h-p-q} = \binom{n-i-j+k}{h-p-q}$ ways.

Thus in n independent trials, the number of possible cases in which A appears i times, B appears j times and $D \cup E \cup F$ appears h times is

$$\binom{n}{k}\binom{n-k}{p}\binom{n-k-p}{i-k-p}\binom{n-i}{q}\binom{n-i-q}{j-k-q}\binom{n-i-j+k}{h-p-q}$$
$$= \frac{n!}{k!p!(i-k-p)!q!(j-k-q)!(h-p-q)!}\frac{1}{(n-i-j+k-h+p+q)!}$$

and each case has equal probability,

$$P(AB)^k P(D)^p [P(AB^c) - P(D)]^{i-k-p} P(E)^q [P(A^cB) - P(E)]^{j-k-q} P(F)^{h-p-q} [P(A^cB^c) - P(F)]^{n-i-j+k-h+p+q}.$$

It is clear that

$$a = \max(0, i + j - n); \qquad b = \min(i, j); \quad i, j, h = 0, 1, \ldots, n. \quad \square$$

Theorem 4.1. *In the fourfold sampling scheme, let $D \subset AB^c$, $E \subset A^cB$, $F \subset A^cB^c$ and ξ, η, ζ be the number of occurrences of the events A, B, $AB \cup D \cup E \cup F$ in n independent trials, respectively (case* (d1) *in Definition 1). Then, the joint probability mass function of ξ, η and ζ is*

$$\begin{aligned} P_{4.1}(i, j, h) &\equiv P\{\xi = i, \eta = j, \zeta = h\} \\ &= \sum_{k=a}^{b}\sum_{p=0}^{i-k}\sum_{q=0}^{j-k} C_{4.1}(n; k, p, q, h, i, j) P(AB)^k \times P(D)^p [P(AB^c) - P(D)]^{i-k-p} P(E)^q \\ &\quad \times [P(A^cB) - P(E)]^{j-k-q} \times P(F)^{h-p-q-k} [P(A^cB^c) - P(F)]^{n-i-j+2k-h+p+q}, \end{aligned} \tag{7}$$

where

$$C_{4.1}(n; k, p, q, h, i, j) = \frac{n!}{k!p!(i-k-p)!q!(j-k-q)!(h-p-q-k)!}\frac{1}{(n-i-j+2k-h+p+q)!};$$
$$a = \max(0, i + j - n); \qquad b = \min(i, j); \quad i, j, h = 0, 1, \ldots, n.$$

Proof. The proof of this theorem is similar to the proof of Theorem 4. $\square$

2. Conditional distributions of bivariate order statistics

Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with distribution functions $F_X(x)$ and $F_Y(y)$, respectively. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a bivariate sample with joint distribution function $F(x, y)$. Additionally, let $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$, $Y_{1:n} \le Y_{2:n} \le \cdots \le Y_{n:n}$ be the corresponding marginal order statistics with distribution functions

$$F_X^{r:n}(x) = P\{X_{r:n} \le x\} = \sum_{i=r}^{n}\binom{n}{i} F(x)^i [1 - F(x)]^{n-i},$$
$$F_Y^{s:n}(y) = P\{Y_{s:n} \le y\} = \sum_{j=s}^{n}\binom{n}{j} F(y)^j [1 - F(y)]^{n-j}.$$

The joint distribution function of $X_{r:n}$ and $Y_{s:n}$ can be obtained easily from the bivariate binomial distribution if one considers the fourfold model with $A = \{X_i \le x\}$ and $B = \{Y_i \le y\}$. Then, $P(AB) = P\{X_i \le x, Y_i \le y\} = \pi_{11}$, $P(AB^c) = P\{X_i \le x, Y_i >$

$y\} = \pi_{12}, P(A^cB) = P\{X_i > x, Y_i \le y\} = \pi_{21}$, and $P(A^cB^c) = P\{X_i > x, Y_i > y\} = \pi_{22}$. If ξ and η are the number of occurrences of events A and B in n independent trials of the fourfold experiment, respectively, then it is clear that

$$P\{X_{r:n} \le x, Y_{s:n} \le y\} = \sum_{i=r}^{n}\sum_{j=s}^{n} P\{\xi = i, \eta = j\}$$
$$= \sum_{i=r}^{n}\sum_{j=s}^{n}\sum_{k=a}^{b} \frac{n!}{k!(i-k)!(j-k)!(n-i-j+k)!}\pi_{11}^{k}\pi_{12}^{i-k}\pi_{21}^{j-k}\pi_{22}^{n-i-j+k},$$

where

$$\pi_{11} = F(x, y),$$
$$\pi_{12} = F_X(x) - F(x, y),$$
$$\pi_{21} = F_Y(y) - F(x, y),$$
$$\pi_{22} = 1 - F_X(x) - F_Y(y) + F(x, y),$$

and $a = \max(0, i + j - n), b = \min(i, j)$ (see [21]).

Now, we are interested in the conditional joint distribution of bivariate order statistics under the condition that h of the bivariate observations $(X_i, Y_i),\ i = 1, 2, \ldots, n$ are truncated and belong to the set

$$\mathbf{B}_{uv} = \{(t, s) \in \mathbb{R}^2 : t \le u, s \le v\}, \quad (u, v) \in \mathbb{R}^2.$$

Lemma 1. *Let (X, Y) be a bivariate random vector with joint distribution function $F(x, y)$ and $(X_1, Y_1), \ldots, (X_n, Y_n)$ be independent copies of (X, Y). If $(X_{r:n}, Y_{s:n})$, $r, s = 1, 2, \ldots, n$ is the vector of bivariate order statistics and $\mathbf{B}$ is any Borel set on $\mathbb{R}^2$, then*

$$F_{r,s:n}(x, y \mid u, v) \equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1, Y_1), \ldots, (X_n, Y_n) \text{ belong to } \mathbf{B}\}$$
$$= \frac{1}{\binom{n}{h} P\{(X, Y) \in \mathbf{B}\}^h P\{(X, Y) \in \mathbf{B}^c\}^{n-h}}$$
$$\times \sum_{i=r}^{n}\sum_{j=s}^{n} P\{\text{exactly } i \text{ of } X's \le x,\ \text{exactly } j \text{ of } Y's \le y, \text{exactly } h \text{ of } (X_i, Y_i)'s \in \mathbf{B}\}, \quad (8)$$

where $\mathbf{B}^c = \mathbb{R}^2 \setminus \mathbf{B}$ is the complement of $\mathbf{B}$.

Proof. From the conditional probability formula, one has

$$P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1, Y_1), \ldots, (X_n, Y_n) \text{ belong to } \mathbf{B}\}$$
$$= \frac{P\{X_{r:n} \le x, Y_{s:n} \le y, h \text{ of } (X_1, Y_1), \ldots, (X_n, Y_n) \text{ belong to } \mathbf{B}\}}{P\{h \text{ of } (X_1, Y_1), \ldots, (X_n, Y_n) \text{ belong to } \mathbf{B}\}}. \quad (9)$$

Because the random vectors $(X_i, Y_i),\ i = 1, 2, \ldots, n$ are assumed to be independent and identically distributed, then from the binomial distribution, one has

$$P\{h \text{ of } (X_1, Y_1), \ldots, (X_n, Y_n) \text{ belong to } \mathbf{B}\} = \binom{n}{h} P\{(X, Y) \in \mathbf{B}\}^h P\{(X, Y) \in \mathbf{B}^c\}^{n-h}. \quad (10)$$

Now, (10) and (9) imply (8). Thus, the lemma is proved. □

For deriving the conditional distribution function of bivariate order statistics $F_{r,s:n}(x, y \mid u, v)$, we consider the following four possible cases:

Case a: $u \le x, v \le y$.
Case b: $u \le x, v > y$.
Case c: $u > x, v \le y$.
Case d: $u > x, v > y$.

Description of Case a. *If $u \le x$, $v \le y$, then we denote $A = \{X_i \le x\}, B = \{Y_i \le y\}$ and $C = \{X_i \le u, Y_i \le v\}$. Let ξ be the number of observations $(X_i, Y_i),\ i = 1, 2, \ldots, n$, for which $X_i \le x$, η be the number of observations for which $Y_i \le y$ and ζ be the number of observations for which $X_i \le u$ and $Y_i \le v$. It is clear that $C \subset AB$ and ξ, η, ζ are the number of observations in n independent trials of the fourfold experiment of the events A, B and C, respectively, as in case* (a) *of Definition 1. We have*

$$P(C) = P\{X \le u, Y \le v\} = F(u, v), \tag{11}$$

$$\begin{aligned} P(AB) - P(C) &= P\{X \le x, Y \le y\} - P\{X \le u, Y \le v\} \\ &= F(x, y) - F(u, v), \end{aligned} \tag{12}$$

$$P(AB^c) = P\{X \le x, Y > y\} = F_X(x) - F(x, y), \tag{13}$$

$$P(A^cB) = P\{X > x, Y \le y\} = F_Y(y) - F(x, y), \tag{14}$$

$$P(A^cB^c) = P\{X > x, Y > y\} = \bar{F}(x, y). \tag{15}$$

Theorem 1a. *If $u \le x, v \le y$, then*

$$\begin{aligned} F^{(1)}_{r,s:n}(x, y \mid u, v) &\equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\} \\ &= \frac{1}{\binom{n}{h} F(u, v)^h [1 - F(u, v)]^{n-h}} \times \sum_{i=r}^{n} \sum_{j=s}^{n} \sum_{k=a}^{b} C_1(n; h, k, i, j) F(u, v)^h \\ &\quad \times [F(x, y) - F(u, v)]^{k-h} [F_X(x) - F(x, y)]^{i-k} [F_Y(y) - F(x, y)]^{j-k} \bar{F}(x, y)^{n-i-j+k}, \end{aligned} \tag{16}$$

$h = 0, 1, \dots, \min(r, s)$ *and*

$F^{(1)}_{r,s:n}(x, y \mid u, v) = 0$ *if* $\min(r, s) < h \le n$,

where

$$C_1(n; h, k, i, j) = \frac{n!}{h!(k-h)!(i-k)!(j-k)!(n-i-j+k)!};$$

$$a = \max(0, i + j - n); \qquad b = \min(i, j).$$

Proof. Because $P\{(X, Y) \in \mathbf{B}_{uv}\} = F(u, v)$ and $P\{(X, Y) \in \mathbf{B}^c_{uv}\} = 1 - F(u, v)$, from Lemma 1, we have

$$\begin{aligned} F^{(1)}_{r,s:n}(x, y \mid u, v) &\equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\} \\ &= \frac{1}{\binom{n}{h} F(u, v)^h [1 - F(u, v)]^{n-h}} \sum_{i=r}^{n} \sum_{j=s}^{n} P\{\xi = i, \eta = j, \zeta = h\}. \end{aligned}$$

Now, (16) easily follows from Theorem 1, from the Description of Case a and the equalities (11)–(15). For $i = r, j = s$, and $h = \min(r, s)$, the probability

$$P\{\xi = i, \eta = j, \zeta = h\}$$

does not vanish. For $i = r + 1, j = s, r < s$, and $h = s + 1$, this probability vanishes because $C \subset AB$ and the number of occurrences of C cannot exceed the number of occurrences of $AB(\{\xi = i, \eta = j\}$ implies that the number of occurrences of AB is $\min(i, j))$. Therefore, for the values of $(i, j) = (r, s), (r + 1, s), (r, s + 1), \dots, (n, n)$, the value of h will vary from 0 to $\min(r, s)$. □

Description of Case b. *If $u \le x, v > y$, then we denote $A = \{X_i \le x\}, B = \{Y_i \le y\}, C = \{X_i \le u, Y_i \le y\}$ and $D = \{X_i \le u, y < Y_i \le v\}$. Let ξ be the number of observations (X_i, Y_i), $i = 1, 2, \dots, n$, for which $X_i \le x$, η be the number of observations for which $Y_i \le y$ and ζ be the number of observations for which $X_i \le u$ and $Y_i \le v$. It is clear that $C \subset AB$, $D \subset AB^c$ and ξ, η, ζ are the number of observations in n independent trials of the fourfold experiment of the events A, B and $C \cup D$, respectively, as in case* (b) *of Definition 1. We have*

$$P(C) = P\{X \le u, Y \le y\} = F(u, y), \tag{17}$$

$$\begin{aligned} P(AB) - P(C) &= P\{X \le x, Y \le y\} - P\{X \le u, Y \le y\} \\ &= F(x, y) - F(u, y), \end{aligned} \tag{18}$$

$$P(D) = P\{X \le u, y < Y \le v\} = F(u, v) - F(u, y), \tag{19}$$

$$\begin{aligned} P(AB^c) - P(D) &= P\{X \le x, Y > y\} - P\{X \le u, y < Y \le v\} \\ &= F_X(x) - F(x, y) - F(u, v) + F(u, y), \end{aligned} \tag{20}$$

$$P(A^cB) = P\{X > x, Y \le y\} = F_Y(y) - F(x, y), \tag{21}$$

$$P(A^cB^c) = P\{X > x, Y > y\} = \bar{F}(x, y). \tag{22}$$

Theorem 2a. *If $u \le x$, $v > y$, then*

$$
\begin{aligned}
F^{(2)}_{r,s:n}(x,y \mid u,v) &\equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1,Y_1),\ldots,(X_n,Y_n) \text{ belong to } \mathbf{B}_{uv}\} \\
&= \frac{1}{\binom{n}{h} F(u,v)^h [1-F(u,v)]^{n-h}} \sum_{i=r}^{n}\sum_{j=s}^{n}\sum_{k=a}^{b}\sum_{l=0}^{h} C_2(n;k,l,h,i,j) F(u,y)^l \\
&\quad \times [F(x,y)-F(u,y)]^{k-l}[F(u,v)-F(u,y)]^{h-l} \\
&\quad \times [F_X(x)-F(x,y)-F(u,v)+F(u,y)]^{i-k-h+l}[F_Y(y)-F(x,y)]^{j-k}\bar{F}(x,y)^{n-i-j+k}, \qquad (23)
\end{aligned}
$$

$h = 0,1,\ldots,r$ *and*

$F^{(2)}_{r,s:n}(x,y \mid u,v) = 0$ *if* $r < h \le n$,

where

$$
C_2(n;k,l,h,i,j) = \frac{n!}{l!(k-l)!(h-l)!(i-k-h+l)!(j-k)!(n-i-j+k)!};
$$

$$
a = \max(0, i+j-n); \qquad b = \min(i,j).
$$

Proof. Similar to the proof of Theorem 1a, the proof of this theorem easily follows from Lemma 1, Theorem 2, Definition 1(b), Description of Case b, and equalities (17)–(22). □

Description of Case c. *If $u > x$, $v \le y$, then we denote $A = \{X_i \le x\}$, $B = \{Y_i \le y\}$, $C = \{X_i \le x, Y_i \le v\}$ and $E = \{x < X_i \le u, Y_i \le v\}$. Let ξ be the number of observations (X_i, Y_i), $i = 1,2,\ldots,n$, for which $X_i \le x$, η be the number of observations for which $Y_i \le y$ and ζ be the number of observations for which $X_i \le u$ and $Y_i \le v$. It is clear that $C \subset AB$, $E \subset A^cB$ and ξ,η,ζ are the number of observations in n independent trials of the fourfold experiment of the events A, B and $C \cup E$, respectively, as in case* (c) *of Definition* 1. *We have*

$$
P(C) = P\{X \le x, Y \le v\} = F(x,v), \qquad (24)
$$

$$
\begin{aligned}
P(AB) - P(C) &= P\{X \le x, Y \le y\} - P\{X \le x, Y \le v\} \\
&= F(x,y) - F(x,v), \qquad (25)
\end{aligned}
$$

$$
P(AB^c) = P\{X \le x, Y > y\} = F_X(x) - F(x,y), \qquad (26)
$$

$$
P(E) = P\{x < X \le u, Y \le v\} = F(u,v) - F(x,v), \qquad (27)
$$

$$
\begin{aligned}
P(A^cB) - P(E) &= P\{X > x, Y \le y\} - P\{x < X \le u, Y \le v\} \\
&= F_Y(y) - F(x,y) - F(u,v) + F(x,v), \qquad (28)
\end{aligned}
$$

$$
P(A^cB^c) = P\{X > x, Y > y\} = \bar{F}(x,y). \qquad (29)
$$

Theorem 3a. *If $u > x$, $v \le y$, then*

$$
\begin{aligned}
F^{(3)}_{r,s:n}(x,y \mid u,v) &\equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1,Y_1),\ldots,(X_n,Y_n) \text{ belong to } \mathbf{B}_{uv}\} \\
&= \frac{1}{\binom{n}{h} F(u,v)^h [1-F(u,v)]^{n-h}} \sum_{i=r}^{n}\sum_{j=s}^{n}\sum_{k=a}^{b}\sum_{l=0}^{h} C_3(n;k,l,h,i,j) F(x,v)^l \\
&\quad \times [F(x,y)-F(x,v)]^{k-l}[F_X(x)-F(x,y)]^{i-k}[F(u,v)-F(x,v)]^{h-l} \\
&\quad \times [F_Y(y)-F(x,y)-F(u,v)+F(x,v)]^{j-k-h+l}\bar{F}(x,y)^{n-i-j+k}, \qquad (30)
\end{aligned}
$$

$h = 0,1,\ldots,s$ *and*

$F^{(3)}_{r,s:n}(x,y \mid u,v) = 0$ *if* $s < h \le n$,

where

$$
C_3(n;k,l,h,i,j) = \frac{n!}{l!(k-l)!(i-k)!(h-l)!(j-k-h+l)!(n-i-j+k)!};
$$

$$
a = \max(0, i+j-n); \qquad b = \min(i,j).
$$

Proof. Similar to the proof of Theorem 2a, the proof of this theorem easily follows from Lemma 1, Theorem 3, Definition 1(c), Description of Case c, and the equalities (24)–(29). □

Description of Case d. *If $u > x$, $v > y$, then we denote $A = \{X_i \le x\}$, $B = \{Y_i \le y\}$, $D = \{X_i \le x, y < Y_i \le v\}$, $E = \{x < X_i \le u, Y_i \le y\}$ and $F = \{x < X_i \le u, y < Y_i \le v\}$. Let ξ be the number of observations (X_i, Y_i), $i = 1,2,\ldots,n$, for which $X_i \le x$, η be the number of observations for which $Y_i \le y$ and ζ be the number of observations for which $X_i \le u$ and $Y_i \le v$.*

It is clear that $D \subset AB^c$, $E \subset A^cB$, $F \subset A^cB^c$ *and* ξ, η, ζ *are the number of observations in* n *independent trials of the fourfold experiment of the events* A, B *and* $AB \cup D \cup E \cup F$, *respectively, as in case* (d1) *of Definition 1. We have*

$$P(AB) = P\{X \le x, Y \le y\} = F(x, y), \tag{31}$$

$$\begin{aligned} P(D) &= P\{X \le x, y < Y \le v\} \\ &= F(x, v) - F(x, y), \end{aligned} \tag{32}$$

$$\begin{aligned} P(AB^c) - P(D) &= P\{X \le x, Y > y\} - P\{X \le x, y < Y \le v\} \\ &= F_X(x) - F(x, v), \end{aligned} \tag{33}$$

$$P(E) = P\{x < X \le u, Y \le y\} = F(u, y) - F(x, y), \tag{34}$$

$$\begin{aligned} P(A^cB) - P(E) &= P\{X > x, Y \le y\} - P\{x < X \le u, Y \le y\} \\ &= F_Y(y) - F(u, y), \end{aligned} \tag{35}$$

$$\begin{aligned} P(F) &= P\{x < X \le u, y < Y \le v\} \\ &= F(u, v) - F(x, v) - F(u, y) + F(x, y), \end{aligned} \tag{36}$$

$$\begin{aligned} P(A^cB^c) - P(F) &= P\{X > x, Y > y\} - P\{x < X \le u, y < Y \le v\} \\ &= 1 - F_X(x) - F_Y(y) - F(u, v) + F(x, v) + F(u, y). \end{aligned} \tag{37}$$

Theorem 4.1a. *If* $u > x$, $v > y$, *then*

$$\begin{aligned} F^{(4.1)}_{r,s:n}(x, y \mid u, v) &\equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\} \\ &= \frac{1}{\binom{n}{h} F(u, v)^h [1 - F(u, v)]^{n-h}} \sum_{i=r}^{n} \sum_{j=s}^{n} \sum_{k=a}^{b} \sum_{p=0}^{i-k} \sum_{q=0}^{j-k} C_{4.1}(n; k, p, q, h, i, j) F(x, y)^k \\ &\quad \times [F(x, v) - F(x, y)]^p [F_X(x) - F(x, v)]^{i-k-p} [F(u, y) - F(x, y)]^q \\ &\quad \times [F_Y(y) - F(u, y)]^{j-k-q} [F(u, v) - F(x, v) - F(u, y) + F(x, y)]^{h-p-q-k} \\ &\quad \times [1 - F_X(x) - F_Y(y) - F(u, v) + F(x, v) + F(u, y)]^{n-i-j+2k-h+p+q}, \end{aligned} \tag{38}$$

$h = 0, \dots, n,$

where

$$C_{4.1}(n; k, p, q, h, i, j) = \frac{n!}{k!p!(i-k-p)!q!(j-k-q)!(h-p-q-k)!} \frac{1}{(n-i-j+2k-h+p+q)!};$$

$$a = \max(0, i + j - n); \qquad b = \min(i, j).$$

Proof. Using Lemma 1, Definition 1(d1), and Description of Case d, one has

$$F^{(4.1)}_{r,s:n}(x, y \mid u, v) = \frac{1}{\binom{n}{h} F(u, v)^h [1 - F(u, v)]^{n-h}} \sum_{i=r}^{n} \sum_{j=s}^{n} P\{\xi = i, \eta = j, \zeta = h\}.$$

Using Theorem 4.1 and equalities (31)–(37), we complete the proof. □

Finally, using the results of Theorems 1a–4.1a, the conditional distribution of bivariate order statistics is presented in the following theorem:

Theorem 5. *Let* (X, Y) *be a bivariate random vector with joint distribution function* $F(x, y)$ *and* $(X_1, Y_1), \dots, (X_n, Y_n)$ *be independent copies of* (X, Y). *If* $(X_{r:n}, Y_{s:n})$, $r, s = 1, 2, \dots, n$, *is the vector of bivariate order statistics and* $\mathbf{B}_{uv} = \{(t, s) \in \mathbb{R}^2 : t \le u, s \le v\}$, $(u, v) \in \mathbb{R}^2$, *then*

$$\begin{aligned} F_{r,s:n}(x, y \mid u, v) &\equiv P\{X_{r:n} \le x, Y_{s:n} \le y \mid h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\} \\ &= \begin{cases} F^{(1)}_{r,s:n}(x, y \mid u, v) & \text{if } u \le x, v \le y, \\ F^{(2)}_{r,s:n}(x, y \mid u, v) & \text{if } u \le x, v > y, \\ F^{(3)}_{r,s:n}(x, y \mid u, v) & \text{if } u > x, v \le y, \\ F^{(4.1)}_{r,s:n}(x, y \mid u, v) & \text{if } u > x, v > y, \end{cases} \end{aligned}$$

$h = 0, 1, \dots, \min(r, s).$

Remark 2. One can verify the accuracy of the results presented in Theorems 1a–4.1a. Here, we present a different method for deriving the conditional distributions of bivariate order statistics using the properties of extreme order statistics $(X_{n:n}, Y_{n:n})$

as follows: Consider

$$F_{n,n:n}(x, y \mid u, v) = P\{X_{n:n} \le x, Y_{n:n} \le y \mid h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{belong to } \mathbf{B}_{uv}\}$$
$$= \frac{1}{\binom{n}{h} F(u,v)^h [1 - F(u,v)]^{n-h}} \times P\{X_{n:n} \le x, Y_{n:n} \le y, h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{belong to } \mathbf{B}_{uv}\}. \quad (39)$$

Because $X_{n:n} \le x$ implies that all $X's$ are less than or equal to x, we can write

$$P\{X_{n:n} \le x, Y_{n:n} \le y, h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\}$$
$$= \sum_{j_1, j_2, \dots, j_n}^{n} P\{X_{n:n} \le x, Y_{n:n} \le y, (X_{j_1}, Y_{j_1}) \in \mathbf{B}_{uv}, \dots, (X_{j_h}, Y_{j_h}) \in \mathbf{B}_{uv},$$
$$(X_{j_{h+1}}, Y_{j_{h+1}}) \in \mathbf{B}^c_{uv}, \dots, (X_{j_n}, Y_{j_n}) \in \mathbf{B}^c_{uv}\}$$
$$= \sum_{j_1, j_2, \dots, j_n}^{n} P\{X_{n:n} \le x, Y_{n:n} \le y, (X_1, Y_1) \in \mathbf{B}_{uv}, \dots, (X_h, Y_h) \in \mathbf{B}_{uv}$$
$$(X_{h+1}, Y_{h+1}) \in \mathbf{B}^c_{uv}, \dots, (X_n, Y_n) \in \mathbf{B}^c_{uv}\} = \binom{n}{h} P\{X_{n:n} \le x, Y_{n:n} \le y, (X_1, Y_1) \in \mathbf{B}_{uv}, \dots, (X_h, Y_h) \in \mathbf{B}_{uv},$$
$$(X_{h+1}, Y_{h+1}) \in \mathbf{B}^c_{uv}, \dots, (X_n, Y_n) \in \mathbf{B}^c_{uv}\} = \binom{n}{h} P\{X_1 \le x, \dots, X_n \le x, Y_1 \le y, \dots, Y_n \le y,$$
$$(X_1, Y_1) \in \mathbf{B}_{uv}, \dots, (X_h, Y_h) \in \mathbf{B}_{uv}, (X_{h+1}, Y_{h+1}) \in \mathbf{B}^c_{uv}, \dots, (X_n, Y_n) \in \mathbf{B}^c_{uv}\}$$
$$= \binom{n}{h} P\{X \le x, Y \le y, (X, Y) \in \mathbf{B}_{uv}\}^h P\{X \le x, Y \le y, (X, Y) \in \mathbf{B}^c_{uv}\}^{n-h}$$
$$= \binom{n}{h} P\{X \le x, Y \le y, X \le u, Y \le v\}^h$$
$$\times P\{X \le x, Y \le y, (X \le u, Y > v \cup X > u, Y \le v \cup X > u, Y > v)\}^{n-h}$$
$$= \binom{n}{h} [P\{X \le \min(x, u), Y \le \min(y, v)\}]^h \times [P\{X \le \min(x, u), v < Y \le y\} + P\{u < X \le x, Y \le \min(y, v)\}$$
$$+ P\{u < X \le x, v < Y \le y\}]^{n-h}. \quad (40)$$

Therefore,

$$F_{n,n:n}(x, y \mid u, v) = \frac{1}{F(u,v)^h [1 - F(u,v)]^{n-h}} [F(\min(x, u), \min(y, v))]^h$$
$$\times [F(\min(x, u), y) - F(\min(x, u), v) + F(x, \min(y, v)) - F(u, \min(y, v))$$
$$+ F(u, v) - F(u, y) - F(x, v) + F(x, y)]^{n-h}. \quad (41)$$

If $u \le x$ and $v \le y$, then we obtain

$$P\{X_{n:n} \le x, Y_{n:n} \le y, h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\}$$
$$= \binom{n}{h} [P\{X \le u, Y \le v\}]^h [P\{X \le u, v < Y \le y\} + P\{u < X \le x, Y \le v\}$$
$$+ P\{u < X \le x, v < Y \le y\}]^{n-h} = \binom{n}{h} F(u, v)^h [F(u, y) - F(u, v) + F(x, v) - F(u, v)$$
$$+ F(u, v) - F(u, y) - F(x, v) + F(x, y)]^{n-h} = \binom{n}{h} F(u, v)^h [F(x, y) - F(u, v)]^{n-h}. \quad (42)$$

Thus, taking into account (42) in (39), we obtain

$$F_{n,n:n}(x, y \mid u, v) = P\{X_{n:n} \le x, Y_{n:n} \le y \mid h \text{ of } (X_1, Y_1), \dots, (X_n, Y_n) \text{ belong to } \mathbf{B}_{uv}\}$$
$$= \frac{[F(x, y) - F(u, v)]^{n-h}}{[1 - F(u, v)]^{n-h}}. \quad (43)$$

Now, let $r = s = n$ in Theorem 1a. Then, it can be easily verified that $F_{n,n:n}(x, y \mid u, v)$ in Theorem 1a equals (43).

Example with graph. Let $F(x, y) = F_X(x)F_Y(y)\{1+\alpha(1-F_X(x))(1-F_Y(y))\}$ be the Farlie–Gumbel–Morgenstern distribution and $F_X(x) = x$, $F_Y(y) = y$, $0 \le x, y \le 1$. This class of distributions has a simple analytical form and is suitable for calculations. Below, we provide a graph of the conditional distribution of bivariate order statistics given in Theorem 5. The graph is drawn using Wolfram Mathematica 7 (see Fig. 1).

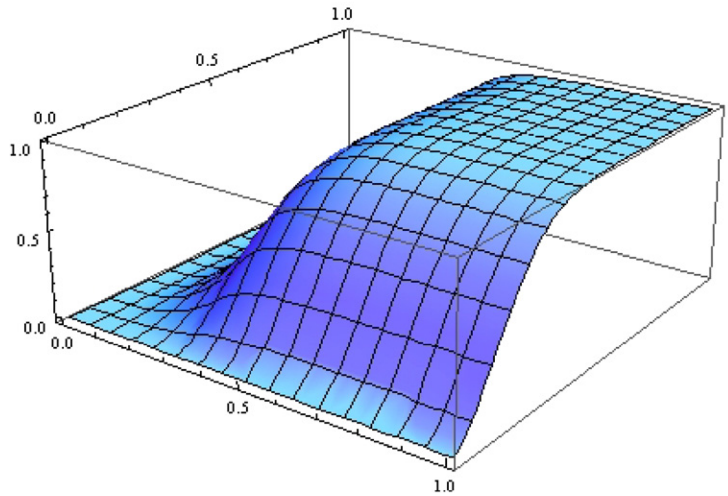

Fig. 1. Graph of $F_{r,s:n}(x, y \mid u, v)$, $n = 10, u = 0.3, v = 0.6, r = 3, s = 2, h = 2, \alpha = 1$.

3. Conclusions

In this paper, we consider novel modifications of bivariate binomial distributions and obtain new trivariate discrete distributions. These distributions are an important class of distributions that are used to derive conditional distributions of bivariate order statistics constructed from a bivariate random sample under the condition that a certain number of observations fall in the given threshold set. The novel trivariate discrete distributions are of interest for distribution theory. The probability generating functions of these distributions are also derived and presented in the Appendix. The conditional distributions of bivariate order statistics presented in Section 2 can be applied widely in many fields of probability and statistics. Note that bivariate order statistics are also important for the construction of new bivariate distributions with high correlation. For example, Baker's-type distributions are constructed on the basis of distributions of bivariate order statistics and attract significant interest in the statistical literature: See, e.g., Bairamov and Bayramoglu [22] and Huang et al. [23]. The findings of Theorem 5 in Section 2 can be used for constructing novel modifications of Baker's-type distributions with high correlation. The results presented in the paper can also be applied widely for reliability analysis of complex systems and studying the dependence among financial markets in crises and other extreme situations.

Acknowledgements

We thank the editor and anonymous reviewers for valuable comments which resulted in improvement of the presentation of this paper.

Appendix

A.1. Probability generating functions

The probability generating function (p.g.f.) of the bivariate binomial distribution (1) is $\Phi(t, s) = (\pi_{11}ts + \pi_{12}t + \pi_{21}s + \pi_{22})^n$. Below, we provide the p.g.f.'s of the trivariate distributions given in Theorems 1–4.1.

Lemma A.1. *Consider the fourfold sampling scheme given in case* (a) *in Definition 1. Then, the joint probability generating function of the random vector (ξ, η, ζ) with probability mass function (p.m.f.) $P_1(i, j, h)$ in (2) in Theorem 1 is*

$$\Phi_1(t, s, z) = (\alpha_1 tsz + \alpha_2 ts + \alpha_3 t + \alpha_4 s + \alpha_5)^n, \tag{44}$$

where

$$\alpha_1 = P(C), \qquad \alpha_2 = P(AB) - P(C), \qquad \alpha_3 = P(AB^c), \qquad \alpha_4 = P(A^cB) \quad \text{and} \quad \alpha_5 = P(A^cB^c).$$

Proof. To derive the joint probability generating functions, let us write

$$\gamma_1^r = \begin{cases} 1 & \text{if in the } r\text{th trial } A \text{ appears,} \\ 0 & \text{otherwise,} \end{cases} \qquad \gamma_2^r = \begin{cases} 1 & \text{if in the } r\text{th trial } B \text{ appears,} \\ 0 & \text{otherwise,} \end{cases}$$

$$\gamma_3^r = \begin{cases} 1 & \text{if in the } r\text{th trial } C \text{ appears,} \\ 0 & \text{otherwise,} \end{cases} \qquad r = 1, 2, \ldots, n.$$

It is clear that $\xi = \sum_{r=1}^{n} \gamma_1^r$, $\eta = \sum_{r=1}^{n} \gamma_2^r$ and $\zeta = \sum_{r=1}^{n} \gamma_3^r$. Because the trials are independent, the p.g.f. of the random vector (ξ, η, ζ) is

$$\Phi(t, s, z) = \left(\sum_{x_1, x_2, x_3 = 0}^{1} t^{x_1} s^{x_2} z^{x_3} q_{x_1, x_2, x_3} \right)^n, \tag{45}$$

where

$$q_{x_1,x_2,x_3} = P\{\gamma_1^r = x_1, \gamma_2^r = x_2, \gamma_3^r = x_3\}; \quad x_1, x_2, x_3 = 0, 1.$$

We have

$$\begin{aligned} q_{1,1,1} &= P(ABC) = P(C), \\ q_{1,1,0} &= P(ABC^c) = P(AB) - P(C), \\ q_{1,0,1} &= P(AB^cC) = 0, \\ q_{0,1,1} &= P(A^cBC) = 0, \\ q_{0,0,1} &= P(A^cB^cC) = 0, \\ q_{0,1,0} &= P(A^cBC^c) = P(A^cB), \\ q_{1,0,0} &= P(AB^cC^c) = P(AB^c), \\ q_{0,0,0} &= P(A^cB^cC^c) = P(A^cB^c). \end{aligned}$$

Then, substituting these values in (45) and simplifying, we obtain (44).

The proofs of the following lemmas are similar. □

Lemma A.2. *Consider the fourfold sampling scheme given in case* (b) *in Definition 1. Then, the joint probability generating function of the random vector* (ξ, η, ζ) *with p.m.f.* $P_2(i, j, h)$ *given in (4) in Theorem 2 is*

$$\Phi_2(t, s, z) = (\alpha_1 tsz + \alpha_2 ts + \alpha_3 tz + \alpha_4 t + \alpha_5 s + \alpha_6)^n, \tag{46}$$

where

$$\alpha_1 = P(C), \quad \alpha_2 = P(AB) - P(C), \quad \alpha_3 = P(D), \quad \alpha_4 = P(AB^c) - P(D),$$
$$\alpha_5 = P(A^cB) \quad \text{and} \quad \alpha_6 = P(A^cB^c).$$

Lemma A.3. *Consider the fourfold sampling scheme given in case* (c) *in Definition 1. Then, the joint probability generating function of the random vector* (ξ, η, ζ) *with p.m.f.* $P_3(i, j, h)$ *given in (5) in Theorem 3 is*

$$\Phi_3(t, s, z) = (\alpha_1 tsz + \alpha_2 ts + \alpha_3 sz + \alpha_4 t + \alpha_5 s + \alpha_6)^n, \tag{47}$$

where

$$\alpha_1 = P(C), \quad \alpha_2 = P(AB) - P(C), \quad \alpha_3 = P(E), \quad \alpha_4 = P(AB^c),$$
$$\alpha_5 = P(A^cB) - P(E) \quad \text{and} \quad \alpha_6 = P(A^cB^c).$$

Lemma A.4. *Consider the fourfold sampling scheme given in case* (d) *in Definition 1. Then, the joint probability generating function of the random vector* (ξ, η, ζ) *with p.m.f.* $P_4(i, j, h)$ *given in (6) in Theorem 4 is*

$$\Phi_4(t, s, z) = (\alpha_1 ts + \alpha_2 tz + \alpha_3 sz + \alpha_4 t + \alpha_5 s + \alpha_6 z + \alpha_7)^n, \tag{48}$$

where

$$\alpha_1 = P(AB), \quad \alpha_2 = P(D), \quad \alpha_3 = P(E), \quad \alpha_4 = P(AB^c) - P(D),$$
$$\alpha_5 = P(A^cB) - P(E), \quad \alpha_6 = P(F), \quad \alpha_7 = P(A^cB^c) - P(F).$$

Lemma A4.1. *Consider the fourfold sampling scheme given in case* (d1) *in Definition 1. Then, the joint probability generating function of the random vector* (ξ, η, ζ) *with p.m.f.* $P_{4.1}(i, j, h)$ *given in (7) in Theorem 4.1 is*

$$\Phi_{4.1}(t, s, z) = (\alpha_1 tsz + \alpha_2 tz + \alpha_3 sz + \alpha_4 t + \alpha_5 s + \alpha_6 z + \alpha_7)^n, \tag{49}$$

where

$$\alpha_1 = P(AB), \quad \alpha_2 = P(D), \quad \alpha_3 = P(E), \quad \alpha_4 = P(AB^c) - P(D),$$
$$\alpha_5 = P(A^cB) - P(E), \quad \alpha_6 = P(F), \quad \alpha_7 = P(A^cB^c) - P(F).$$

References

[1] A.C. Aitken, H.T. Gonin, On fourfold sampling with and without replacement, Proceedings of the Royal Society of Edinburgh 55 (1935) 114–125.
[2] A.S. Krishnamoorthy, Multivariate binomial and Poisson distributions, Sankhyā: The Indian Journal of Statistics 11 (2) (1951) 117–124.
[3] M.A. Hamdan, H.A. Al-Bayyati, A note on the bivariate Poisson distribution, The American Statistician 23 (4) (1969) 32–33.
[4] M.A. Hamdan, Canonical expansion of the bivariate binomial distribution with unequal marginal indices, International Statistical Review 40 (1972) 277–280.
[5] M.A. Hamdan, A note on the trinomial distribution, International Statistical Review 43 (1975) 219–220.
[6] D.N. Shanbhag, I.V. Basawa, On a characterization property of the multinomial distribution, Trabajos de Estadistica 25 (1974) 109–112.
[7] D.C. Doss, R.C. Graham, A characterization of multivariate binomial distribution by univariate marginals, Calcutta Statistical Association Bulletin 24 (1975) 93–99.
[8] M.A. Hamdan, D.R. Jensen, A bivariate binomial distribution and some applications, Australian Journal of Statistics 18 (3) (1976) 163–169.
[9] S. Kocherlakota, K. Kocherlakota, Bivariate Discrete Distributions, Marcel Dekker Inc., New York, 1992.
[10] N.L. Johnson, S. Kotz, N. Balakrishnan, Discrete Multivariate Distributions, John Wiley & Sons, New York, 1997.
[11] I. Bairamov, O.E. Gultekin, Discrete distributions connected with bivariate binomial, Hacettepe Journal of Mathematics and Statistics 39 (1) (2010) 109–120.
[12] F. Durante, P. Jaworski, Spatial contagion between financial markets: a copula-based approach, Applied Stochastic Models in Business and Industry 26 (2010) 551–564.
[13] K.H. Bae, G.A. Karolyi, R.M. Stulz, A new approach to measuring financial contagion, Review of Financial Studies 16 (3) (2003) 717–763.
[14] A. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press, USA, 2005.
[15] B.O. Bradley, M.S. Taqqu, How to estimate spatial contagion between financial markets, Finance Letters 3 (1) (2005) 64–76.
[16] J.C. Rodriguez, Measuring financial contagion: a Copula approach, Journal of Empirical Finance 14 (2007) 401–423.
[17] O.L. Gebizlioglu, B. Yagci, Tolerance intervals for quantiles of bivariate risks and risk measurement, Insurance: Mathematics and Economics 42 (3) (2008) 1022–1027.
[18] S.K. Acik, O.L. Gebizlioglu, Risk analysis under progressive type II censoring with binomial claim numbers, Journal of Computational and Applied Mathematics 233 (1) (2009) 61–72.
[19] F. Tank, O.L. Gebizlioglu, A. Apaydin, Determination of dependency parameter in joint distribution of dependent risks by fuzzy approach, Insurance: Mathematics and Economics 38 (1) (2006) 189–194.
[20] I. Bairamov, Reliability and mean residual life of complex systems with two dependent components per element, IEEE Transactions on Reliability (2013) (in press).
[21] H.A. David, Order Statistics, second ed., John Wiley & Sons, New York, 1981.
[22] I. Bairamov, K. Bayramoglu, From the Huang–Kotz FGM distribution to Baker's bivariate distribution, Journal of Multivariate Analysis 113 (2013) 106–115.
[23] J.S. Huang, X. Dou, S. Kuriki, G.D. Lin, Dependence structure of bivariate order statistics with applications to Bayramoglu's distributions, Journal of Multivariate Analysis 114 (2013) 201–208.