TRIVARIATE BINOMIAL DISTRIBUTION AND TRIVARIATE ORDER STATISTICS

ÇAĞIN KESKİN

Master's Thesis

Graduate School

İzmir University of Economics

İzmir

2020

TRIVARIATE BINOMIAL DISTRIBUTION AND TRIVARIATE ORDER STATISTICS

ÇAĞIN KESKİN

A Thesis Submitted to
The Graduate School of İzmir University of Economics
Master's Program in Applied Statistics

İzmir
2020

ABSTRACT

TRIVARIATE BINOMIAL DISTRIBUTION AND TRIVARIATE ORDER STATISTICS

Keskin, Çağın

M.Sc. in Applied Statistics

Advisor: Prof. Dr. İsmihan Bayramoğlu

July, 2020

In this thesis, trivariate binomial equation derived by using bivariate binomial with fourfold scheme. Applying this equation to order statistics, trivariate order statistics distribution was obtained. In the fourth part of the thesis, Gumble copula was used as a special example for trivariate order statistics distribution and the subject was expanded with numerical examples and applications. These new distributions can be used in probability models and the theoretical studies of the field of statistics. Furthermore, different solutions can be developed by integrating into game theory studies in economics.

Keywords: Trivariate binomial distribution, trivariate order statistics, fourfold scheme, gumbel copula, bivariate binomial distribution

ÖZET

ÜÇ DEĞİŞKENLİ BİNOM DAĞILIMI VE ÜÇ DEĞİŞKENLİ SIRA İSTATİSTİKLERİ

Keskin, Çağın

Uygulamalı İstatistik Yüksek Lisans Programı

Tez Danışmanı: Prof. Dr. İsmihan Bayramoğlu

Temmuz, 2020

Bu tezde, dört katlı şema gösterimine sahip iki değişkenli binom dağılımı kullanılarak, üç değişkenli binom denklemi türetilmiştir. Bu denklemi sıra istatistiklerine uygulayarak üç değişkenli sıra istatistiği dağılımı elde edilmiştir. Tezin dördüncü bölümünde, Gumbel kopula, üç değişkenli sıra istatistikleri dağılımında özel örnek olarak kullanıldı, ayrıca konu sayısal örnekler ve uygulamalar ile genişletildi. Bu yeni dağılımlar olasılık modellerinde ve istatistik alanının teorik çalışmalarında kullanılabilir. Ek olarak, iktisatta oyun teorisi çalışmalarına entegre edilerek farklı çözümler geliştirilebilir.

Anahtar Kelimeler: Üç değişkenli binom dağılımı, üç değişkenli sıra istatistikleri, dört katlı şema, gumbel kopula, iki değişkenli binom dağılımı

Dedicated to my family, my friends and all my teachers…

ACKNOWLEDGEMENT

I would like to present my deepest gratitude to my advisor Prof. Dr. İsmihan Bayramoğlu. Since the beginning of the thesis, he has been always helpful and guiding me to successfully complete and present this document. Without my advisor, this thesis and study would not have been. It has been an honor and privilege to be his student and to work with him.

Moreover, I also would like to thank, my dear friends, Hasan Akbaba, Çağatay Albir, and Aslı Egeli for their valuable friendship. With them, I had a lovely and enjoyable university life. They have always motivated me in this long and difficult thesis writing process.

Finally, I would like to endless thank my dear mother Songül Ergün and my father Hasan Keskin for their love and support in all aspects of my life. Without them, I could not have gone so far in my life and strive for more.

In addition, I would like to thank everyone who was not mentioned here but contributed to this challenging, instructive and long journey

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In statistics, the Bernoulli distribution named after the Swedish scientist Jakob Bernoulli (1713) and improved version of this distribution (binomial distribution) can be used in many areas of statistics. The Bernoulli distribution is basically two options of an event with shown probability p and $1\text{-}p$. While Bernoulli distribution occurs a single event, binomial distribution has n independent repetitions of Bernoulli events. Yule (1919) processed the graphical and theoretical methods of the binomial distribution, after that Aitken, and Gonin (1936) demonstrated the bivariate binomial distribution by creating a fourfold table without replacement. Hamdan (1972) showed the bivariate binomial distribution for X_1 and X_2 with unequal samples of n_1 and n_2,, Oluyede (1994) has worked on the bivariate binomial by deriving the equal and unequal margin indexes from the bivariate bernoulli distribution. Hamdan, and Jensen (1976), Kocherlakota (1989) has made its various applications, which are conditional distribution and regression. In addition to that, maximum likelihood estimation has applied to the bivariate binomial (Hamdan, and Martinson, 1971) and the parameters of the distribution (Hamdan, and Nasro, 1986).

Bivariate of binomial, poisson, negative binomial, hypergeometric, geometric, exponential and gamma distributions were derived Marshall, and Olkin (1985) based on the bivariate Bernoulli. Crowder, and Sweeting (1989) implemented the margins of the bivariate binomial distribution by bayesian inference $X \mid n \sim Binomial\ (n,p)$ and $Y \mid X,n \sim Binomial(n,q)$. Polson, and Wasserman (1990) discussed the bivariate binomial distribution using different priorities. Lee (1984) used symmetric bivariate binomial in the health field by dividing patients two cluster sample groups, retinitis pigmentosa and otitis media. On the other hand, Mishra (1996), made the generalized formula of the bivariate binomial distribution with three parameters. Biswas, and Hwang (2002) made a new derivation of bivariate binomial when X and Y random variables have a correlation.

Papageorgiou, and David (1994) obtained the factorial moments, factorial cumulative and conditional distributions of the bivariate binomial distribution and derived the mixture bivariate binomial from the poisson distribution. They (Papageorgiou, and David, 1995) also showed trivariate binomial distribution on probability generating functions and conditional distributions. Chandrasekar, and Balakrishnan (2002) showed the trivariate binomial distribution on the regression equation.

Ordered statistics are also widely used in non-parametric and inference statistics. Although it was used many papers, Wilks was the first systematically addressed order statistics both non-parametric (1942) and parametric (1948) methods. Boland et al. (1996) examined the dependency of bivariate order statistics under different conditions. The moments of bivariate order statistics have developed (Barakat, 1999) and Bairamov, and Eryilmaz (2003) studied bivariate order statistics (X_n, Y_n) with the new order of sample pairs (X_{n+m}, Y_{n+m}). Bairamov, and Gültekin (2010) worked on the different parameter margins of the bivariate binomial and its extensions of 3x3 table. On the other hand, Bairamov, and Kemalbay (2013) examined conditional bivariate order statistics on the improved version of the fourfold scheme. In addition, they (2015) took Bairamov, and Eryilmaz derivation of the new order of sample pairs and applied it in order to bivariate binomial distributions.

In this thesis, trivariate binomial distribution is shown and trivariate order statistics are studied. The thesis proceeds as follows, bivariate binomial distributions and their modifications are given in the second chapter. In the third chapter, bivariate order statistics and how to obtain from bivariate binomial distribution is discussed. The fourth chapter is examined under five subtitles and mainly focused on the trivariate binomial distribution and trivariate order statistics. Afterwards, the illustration of Gumbel copula was given and some numerical results and applications were followed. In the last chapter, scope of the subject and its integration into different areas will be discussed.

CHAPTER 2: BIVARIATE BINOMIAL DISTRIBUTION

Bivariate Bernoulli distribution and a detailed explanation of the bivariate binomial will be discussed in this section (Aitken, and Gonin, 1936). Subsequently, various modifications of bivariate binomial will be illustrated and the size of the subject will be expanded.

2.1 *Bivariate Bernoulli Distribution*

Let $\{\Omega, \mathcal{F}, P\}$ be a probability space and $X(\omega)$ and $Y(\omega)$ be discrete random variables defined in this space where $\omega \in \Omega$

$$X(\omega) = \begin{cases} 1, & if \ \ \omega \in A \\ 0, & if \ \ \omega \in A^C \end{cases}$$

$$Y(\omega) = \begin{cases} 1, & if \ \ \omega \in B \\ 0, & if \ \ \omega \in B^C \end{cases}$$

where $A, B \in \mathcal{F}$ and distribution of probability mass function (pmf) presented as,

$$P\{X = 1\} = p, P\{X = 0\} = 1 - p, \tag{2.1.1}$$

$$P\{Y = 1\} = q, P\{Y = 0\} = 1 - q, \tag{2.1.2}$$

and $0 < p < 1, \ 0 < q < 1$.

Probability mass functions, as expressed above are called the Bernoulli distribution with the parameter p and q. Furthermore, is shown as $X \sim Bernoulli(p)$ and $Y \sim Bernoulli(q)$, respectively. Teugels (1990) made the bivariate Bernoulli representation with parameters p_1 and q_1 in fourfold table. Let $X \sim Bernoulli(p_1)$ and $Y \sim Bernoulli(q_1)$ be bivariate Bernoulli distribution, four occurrences can be happen AB, AB^C, A^CB, A^CB^C with the probabilities π_{11}, π_{12}, π_{21}, π_{22}. In this case, bivariate Bernoulli distribution is defined,

$$P\{X = 1, Y = 1\} = P(AB) = \pi_{11}, \tag{2.1.3}$$
$$P\{X = 1, Y = 0\} = P(AB^C) = \pi_{12},$$
$$P\{X = 0, Y = 1\} = P(A^C B) = \pi_{21},$$
$$P\{X = 0, Y = 0\} = P(A^C B^C) = \pi_{22}.$$

From (2.1.1) and (2.1.2) marginals of the bivariate Bernoulli distribution are

$$P\{X = 1\} = p_1 = \pi_{11} + \pi_{12}\ , P\{X = 0\} = p_2 = \pi_{21} + \pi_{22} \quad (2.1.3)$$

$$P\{Y = 1\} = q_1 = \pi_{11} + \pi_{21}\ , P\{Y = 0\} = q_2 = \pi_{12} + \pi_{22} \quad (2.1.4)$$

where

$$\sum \pi_{lm} = 1, \qquad \text{where } l, m = 1, 2.$$

Table 2.1: Bivariate Bernoulli Distribution in 2x2 Matrix

A | B	B	B^C	$\sum$
A	A B π_{11}	A B^C π_{12}	p_1
A^C	A^C B π_{21}	A^C B^C π_{22}	p_2
$\sum$	q_1	q_2	1

After defining margins and probabilities; expectations, variances, covariance and correlation can be derived as below.

$$E(X) = p_1 \quad E(Y) = q_1 \tag{2.1.4}$$

$$Var(X) = p_1 \times p_2 \quad Var(Y) = q_1 \times q_2 \tag{2.1.5}$$

$$Cov(X,Y) = E(XY) - E(X)\,E(Y) \tag{2.1.6}$$

$$Cov(X,Y) = \pi_{11} - p_1 \times q_1 = \pi_{11}\pi_{22} - \pi_{12}\pi_{21}$$

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)}\,\sqrt{Var(Y)}} \tag{2.1.7}$$

$$\rho = \frac{\pi_{11}\pi_{22} - \pi_{12}\pi_{21}}{\sqrt{(\pi_{11}+\pi_{12})\,(\pi_{21}+\pi_{22})}\;\sqrt{(\pi_{11}+\pi_{21})}\,\sqrt{(\pi_{12}+\pi_{22})}}$$

For more detailed information about the Bernoulli distribution, the probability generating function (Teugels, 1990) and different distributions (Marshall, and Olkin, 1985) are derived from bivariate Bernoulli can be found. In this thesis, only the bivariate Bernoulli, probability table and probability mass function are emphasized for the representation of bivariate binomial distribution.

2.2 *Bivariate Binomial Distribution*

The binomial distribution is obtained from *n* independent repetitions of the Bernoulli distribution and Jakob Bernoulli introduced this distribution. Probability mass function of *k* successes in *n* independent Bernoulli trial is expressed as,

$$P\{X = k\} = \binom{n}{k} p^k \,(1-p)^{n-k} \tag{2.2.1}$$

where $k = 1,2,\dots,n$ and $\in (0,1)$.

The distribution illustrated above is called binomial distribution with the parameter *n, p* and it is shown $X \sim Binomial(n,p)$. Let X and Y be a binomial random variables and shown these random variables as an independent bivariate sample $(X_1,Y_1),(X_2,Y_2),(X_3,Y_3),\dots,(X_n,Y_n)$ with,

$$X(\omega) \ = \begin{cases} 1, & if \ \ \omega \in A \\ 0, & if \ \ \omega \in A^C, \end{cases}$$

$$Y(\omega) \ = \begin{cases} 1, & if \ \ \omega \in B \\ 0, & if \ \ \omega \in B^C. \end{cases}$$

$\{\Omega \, , \mathcal{F} \, , P\}$ be probability space and $X(\omega)$,$Y(\omega)$ discrete variable of this space where $\omega \in \Omega$ and $A \, , B \ \in \ \mathcal{F}$. Aitken, and Gonin (1936) are defined *X* and *Y* binomial random variables with two possible outcomes *A*, A^C and *B*, B^C. They presented bivariate binomial distribution on fourfold scheme without replacement with *AB*, AB^C, A^CB, A^CB^C events and $P(AB) = \pi_{11}, P(AB^C) = \pi_{12}, P(A^CB) = \pi_{21}, P(A^CB^C) = \pi_{22}$ probabilities.

$$\sum \pi_{lm} \ = \ 1 \, , \qquad \text{where } l, m \ = \ 1, 2.$$

$\xi_1 \, , \xi_2 \, , \xi_{11}, \xi_{12}, \xi_{13}$ defined number of occurrences of *n* experiment as below.

$$\xi_1 \ = \ \{ \, number \ of \ occurences \ A \ in \ n \ trials\}$$

$$\xi_2 \ = \ \{number \ of \ occurences \ B \ in \ n \ trials\}$$

$$\xi_{11} \ = \ \{ \, number \ of \ occurences \ AB \ in \ n \ trials\}$$

$$\xi_{12} \ = \ \{ \, number \ of \ occurences \ AB^C \ in \ n \ trials\}$$

$$\xi_{13} \ = \ \{ \, number \ of \ occurences \ A^CB \ in \ n \ trials\}$$

In addition,

$$\xi_1 \ = \ \xi_{11} \ + \ \xi_{12} \ \ , \xi_2 \ = \ \xi_{11} \ + \ \xi_{13} \tag{2.2.2}$$

$$\xi_1 \ = \ \sum_{i=1}^{n} X_i \ \text{and} \ \xi_2 \ = \ \sum_{j=1}^{n} Y_j \, , \ \ i, j \ = 1,2,3 \ldots n \, .$$

Table 2.2: Bivariate Binomial Distribution on Fourfold Scheme

A | B	B	B^C
A	A B π_{11} *h times*	A B^C π_{12} *i-h times*
A^C	A^C B π_{21} *j-h times*	$A^C B^C$ π_{22} *n-i-j+h times*

Consider the situation $\xi_1 = i \text{ and } \xi_2 = j$, while n is the total number of repetitions. Probability mass function of bivariate binomial can be written as follows,

$$P(i,j) = P\{\xi_1 = i, \xi_2 = j\}$$

$$= \sum_{h=max(i+j-n,\ 0)}^{min(i,j)} C(i,j,h,n)\ {\pi_{11}}^{h}\ {\pi_{12}}^{i-h}\ {\pi_{21}}^{j-h}\ {\pi_{22}}^{n-i-j+h} \tag{2.2.3}$$

where

$$C(i,j,k,n) = \binom{n}{h}\binom{n-h}{i-h}\binom{n-i}{j-h}\binom{n-i-j+h}{i-h}$$

Thus, open form of the equation will be,

$$\sum_{h=\mathrm{m}\ (i+j-n,\ 0)}^{\min(i,j)} \frac{n!}{h!\,(i-h)!\,(j-h)!\,(n-i-j+h)!}\ {\pi_{11}}^{h}\ {\pi_{12}}^{i-h}\ {\pi_{21}}^{j-h}\ {\pi_{22}}^{n-i-j+h}$$

and $i,j = 0,1,2,3\ldots n.$ (2.2.4)

Bivariate binomial distribution with four probability situations have shown by many studies (Aitken, and Gonin, 1936; Kocherlakota, 1989). From n experiments, if A and B occurred together h times, they can be figure out $\binom{n}{h}$ ways. Since A has a total of i observation, A and B^C occurred together i-h times and they can be realized $\binom{n-h}{i-h}$ ways. B has a j observations, therefore it can be observed together A^C with j-h times, and they can be figure out $\binom{n-i}{j-h}$ ways. It is clear that, when A and B occurrences known, the outcomes of AB must be appeared upper bound $min(i, j)$ and the lower bound $max(i + j - n,\ 0)$ observations. Hence, the probability of bivariate binomial is the sum of all possible states. If the exact value of h is known, all possibilities of h do not need to be realized and $P\,\{\xi_1 = i\,, \xi_2 = j\,, \xi_{11} = h\,\}$ equals to

$$= \frac{n!}{h!\,(i-h)!\,(j-h)!\,(n-i-j+h)!}\,{\pi_{11}}^{h}\,{\pi_{12}}^{i-h}\,{\pi_{21}}^{j-h}\,{\pi_{22}}^{n-i-j+h}.$$

After bivariate binomial marginals have been identified, expectations, variances and covariance are

$$E(\xi_1) = n\,(\pi_{11} + \pi_{12})\,, \qquad E(\xi_2) = n\,(\pi_{11} + \pi_{21}) \tag{2.2.5}$$

$$\begin{aligned} Var\,(\xi_1) &= n\,(\pi_{11} + \pi_{12})\,(\pi_{21} + \pi_{22}) \\ Var\,(\xi_2) &= n\,(\pi_{11} + \pi_{21})\,(\pi_{12} + \pi_{22}) \end{aligned} \tag{2.2.6}$$

$$Cov(\xi_1, \xi_2) = n\,[\pi_{11}\pi_{22} - \pi_{12}\pi_{21}] \tag{2.2.7}$$

The general lines of the bivariate binomial distribution and the fourfold scheme are emphasized to create the trivariate binomial distribution. More detailed research on bivariate binomial distribution, can be found at Hamdan's works (Hamdan, 1972; Hamdan, and Nasro, 1986; Hamdan, and Jensen, 1976; Hamdan, and Martinson, 1971).

2.3 *Modification of Bivariate Binomial Distribution*

In the previous subsection, bivariate binomial distribution is defined in the fourfold model and processed the margins and probability mass function. After these are described, modifications of the bivariate binomial distribution can be simply derived and expand it for different conditions. Bairamov, and Gültekin (2010) modified the bivariate binomial random variable three possibilities, A_1 , A_2 , A_3 for *X* and B_1 , B_2 , B_3 for *Y*. Thus, there were nine possible results which are $A_1B_1 , A_1B_2 , A_1B_3 , A_2B_1, A_2B_2 , A_2B_3, A_3B_1 , A_3B_2 , A_3B_3$ wit $\pi_{11}, \pi_{12}, \pi_{13}, \pi_{21}, \pi_{22}, \pi_{23}, \pi_{31}, \pi_{32}, \pi_{33}$ probabilities. Furthermore, ξ_1 symbolize the number of times *n* experiment observed in A_1, and $\xi_2 , \xi_{12} , \xi_{21}$ are number of occurrences in *n* trial B_1 , A_1B_2 *and* A_2B_1, respectively.

Assume that $\xi_1 = i , \xi_2 = j , \xi_{12} = r , \xi_{21} = m$. Under these condition, $P(i,j,r,m) = P\,\{\xi_1 = i , \xi_2 = j , \xi_{12} = r , \xi_{13} = m\}$ will be

$$= \sum_{h=max(i+j-n,\ 0)}^{min(i-r,j-m)} C(i,j,h,r,m,n)\ {\pi_{11}}^h {\pi_{12}}^r\ {\pi_{13}}^{i-r-h}\ {\pi_{21}}^m\ {\pi_{31}}^{j-h-m}\ (\pi^c)^{\,n-i-j+h} \tag{2.3.1}$$

where $i = r, \ldots, n-m\ ;\ j = m, \ldots, n-r\ ;\ r = 0, \ldots, n-m\ ;\ m = 0, \ldots, n$

and

$$1 - \pi_{11} - \pi_{12} - \pi_{13} - \pi_{21} - \pi_{31} = \pi^c$$

and

$$C(i,j,h,r,m,n) = \binom{n}{r}\binom{n-r}{h}\binom{n-r-h}{i-h-r}\binom{n-i}{m}\binom{n-i-m}{j-h-m}$$

$$= \frac{n!}{h!\,r!\,m!\,(i-h-r)!\,(j-h-m)!\,(n-i-j+h)!}$$

In n trials, A_1 and B_2 occurred together r times and they can be realized $\binom{n}{r}$ ways. A_1 and B_1 observed together h times, therefore they can be figure out $\binom{n-r}{h}$ ways. Since, A_1 has a total of i observation, A_1 and B_3 occurred together *i-r-h* times and they can be realized $\binom{n-r-h}{i-r-h}$ ways. Moreover, A_2 and B_1 occurred together m times, therefore it can be realized $\binom{n-i}{m}$ ways. B_1 has a j observations, therefore it can be appeared together A_3 with *j-m-h* times, and they can be figure out $\binom{n-i-m}{j-m-h}$ ways. It is obvious that, when A_1 B_2, A_2 B_1, A_1 and B_1 occurrences known, the outcomes of A_1 B_1 must be appeared upper bound $min(i-r, j-m)$ and the lower bound $max(i+j-n,\ 0)$ observations.

Table 2.3: Modification of Bivarite Binomial Distribution on 3x3 Matrix

A | B	B_1	B_2	B_3
A_1	A_1B_1 π_{11} *h times*	A_1B_2 π_{12} *r times*	A_1B_3 π_{13} *i-h-r times*
A_2	A_2B_1 π_{21} *m times*	A_2B_2 π_{22}	A_2B_3 π_{23}
A_3	A_3B_1 π_{31} *j-h-m times*	A_3B_2 π_{32}	A_3B_3 π_{33}

Bairamov, and Kemalbay (2013) made the more challenging bivariate binomial modification by adding subgroup events to three sets of possible outcomes. Let, E, D and F are the subset events of A^CB, AB^C, A^CB^C, respectively. ξ_1, ξ_2 and η be total number of events take place in *A, B*, and $(D \cup F \cup E)$ with *i ,j ,k*. Hence, there are four main events AB, AB^C, A^CB, A^CB^C and three subset events E, D and F with probabilities $\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22}$ of main events and p_1, p_2, p_3 of subset events, respectively.

Under these conditions, $P(i,j,k) = P\{\xi_1 = i, \xi_2 = j, \eta = k\}$ of bivariate binomial is presented,

$$= \sum_{h=a}^{b} \sum_{p=0}^{i-h} \sum_{q=0}^{j-h} C(i,j,h,r,q,k,n)\ \pi_{11}{}^{h}\, p_1{}^{r}\, [\pi_{12} - p_1]^{i-h-r} \tag{2.3.2}$$

$$\times\, p_2{}^{q}\, [\pi_{21} - p_2]^{j-h-q}\, p_3{}^{k-r-q}\, [\pi_{22} - p_3]^{n-i-j+h-k+r+q}$$

where $C(i,j,h,r,q,k,n)$

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

and

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

In n trials, A and B occurred together h times and they can be realized $\binom{n}{h}$ ways. D, which is the subset events of AB^C, observed r times, they can be figure out $\binom{n-h}{r}$ ways. AB^C observed i-h times and they can be realized $\binom{n-h-r}{i-h-r}$ ways. E is the subset event of A^CB and appeared q times, therefore it can be figure out $\binom{n-i}{q}$ ways. A^CB appeared together j-h times and they can be realized $\binom{n-i-q}{j-h-q}$ ways. Since, total number of subset events k, thus F occurred k-p-q times and it can be figure out $\binom{n-i-j+h}{k-p-q}$ ways. It is clear that, when A and B occurrences known, the outcomes of AB must be appeared upper bound $min(i,j)$ and the lower bound $max(i+j-n,\ \ 0)$ observations.

Table 2.4: Bivarite Binomial Distribution with subset events E, D and F

A | B	B	B^C
A	A B π_{11} *h times*	A B^C π_{12} *i-h times* D p_1 *r times*
A^C	A^C B π_{21} *j-h times* E p_2 *q times*	A^C B^C π_{22} *n-i-j+h times* F p_3 *k-r-q times*

Conditional distribution (Bairamov, and Kemalbay, 2013) bayesian statistics (Crowder, and Sweeting, 1989) and priorities (Polson, and Wasserman, 1990) of bivariate binomial are beyond the scope of this thesis, however they are important modifications.

CHAPTER 3: BIVARIATE ORDER STATISTICS

In chapter 3, single-order statistics will be defined and their distribution are obtained from incomplete beta distribution. Moreover, using bivariate binomial equation, bivariate order statistics distribution will be derived and explained.

3.1 Distribution of Single Order Statistics

Let $X_1, X_2, \dots, X_n$ be an independent identically distributed random variable samples from infinite population with cumulative distribution function of *F*. Ordering the $X_1, X_2, \dots, X_n$ correspond to increasing order as follows,

$$X_{(1)} = min(X_1, X_2, \dots, X_n)$$

$$.$$

$$X_{(i)} = i'th\ smallest\ of\ X_1, X_2, \dots, X_n$$

$$.$$

$$X_{(n)} = max(X_1, X_2, \dots, X_n).$$

Thus, $X_{(1)} \leq X_{(2)} \leq, \dots, \leq X_{(n)}$ are dependent random variables and called order statistics of *X*. $F_X(x)$ is cumulative distribution function of each random variable samples *X* and the cumulative distribution function of *r* th order statistic be represented $F_{(r)}(x)$. In addition, n sample size of *r* th ordered statistics can be displayed $X_{(r)}$ or $X_{r:n}$. According to definition, $F_n(x)$ and $F_1(x)$ can be written,

$$P\{X_{(n)} \leq x\} = P\{X_{(1)} \leq x , X_{(2)} \leq x , \dots, X_{(n)} \leq x\}$$

$$F_{max}(x) \equiv F_n(x) = P\{X_1 \leq x\}, P\{X_2 \leq x\}, \dots, P\{X_n \leq x\} = (F_X(x))^n \tag{3.1.1}$$

and

$$F_{min}(x) \equiv F_1(x) = 1 - P\{X_1 > x\} = 1 - P\{X_1 > x, X_2 > x, \dots, X_n > x\}$$

$$= 1 - (1 - P\{X_1 \le x\}), (1 - P\{X_2 \le x\}), \dots, (1 - P\{X_n \le x\})$$

$$= 1 - (1 - F_X(x))^n \qquad (3.1.2)$$

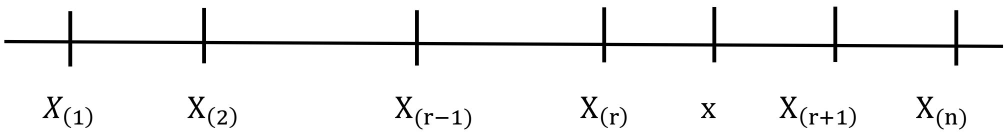

Figure 3.1: Single Order Statistics Illustration

The $F_1(x)$ and $F_n(x)$ is expressed above and to find $F_{(r)}(x)$,

$$F_{(r)}(x) = P\{X_{(r)} \le x\} = P\left\{\bigcup_{i=r}^{n} \{exacly\ i\ of\ X_1, X_2, \dots, X_n\ are\ \le x\}\right\}$$

$$= \sum_{i=r}^{n} P\{exacly\ i\ of\ X_1, X_2, \dots, X_n\ are\ \le x\} = \sum_{i=r}^{n} C_n^i (F_X(x))^i (1 - F_X(x))^{n-i}$$

(3.1.3)

3.1.1 Beta Distribution

Beta distribution is bounded to [0, 1] and it is a distribution derived from the beta function (B). Beta function can be expressed,

$$B(a, \beta) = \int_0^1 x^{a-1}(1-x)^{\beta-1} dx = \frac{\Gamma(a)\,\Gamma(\beta)}{\Gamma(a+\beta)} = \frac{(a-1)!\,(\beta-1)!}{(a+\beta-1)!}.$$

(3.1.1.1)

Let X be a continuous random variable with parameter $0 \le x \le 1$ and $a, \beta > 0$. Then, cdf of Beta distribution will be,

$$Beta_{cdf}(x;a,\beta) = \begin{cases} \frac{1}{B(a,\beta)} \int_0^{F_X(x)} x^{a-1}(1-x)^{\beta-1} dx \\ 0\,, \qquad\qquad otherwise, \end{cases} \tag{3.1.1.2}$$

and probability density function (pdf) of Beta distribution will be,

$$Beta_{pdf}(x;a,\beta) = \begin{cases} \frac{1}{B(a,\beta)} x^{a-1}(1-x)^{\beta-1} \\ 0\,, \qquad\qquad otherwise. \end{cases} \tag{3.1.1.3}$$

It is known that cdf of rth order statistics can be written,

$$\sum_{i=r}^{n} C_n^i (F_X(x))^i (1-F_X(x))^{n-i} = \frac{1}{\beta(r,n-r+1)} \int_0^{F_X(x)} x^i (1-x)^{n-i} dt, \tag{3.1.1.4}$$

which is known incomplete beta distribution and denoted by $I_{F_{(x)}}(r, n-r+1)$.

Furthermore, probability density function of rth order statistics can be derived,

$$P\{r-1\ of\ X_1, X_2, \dots, X_n \le x, one\ is\ in\ (x, x+\Delta x), and\ n-r\ of\ > x+\Delta x\}$$

$$= \binom{r-1}{n} \binom{1}{n-(r-1)} \binom{n-r}{n-(r-1)-1}$$

$$\times \big(F(x)\big)^{r-1} \big(F(\,x+\Delta x) - F(x)\big) \big(1 - F(\,x+\Delta x)\big)^{n-r}$$

$$= \frac{n!}{(r-1)!\,(n-r)!} \big(F(x)\big)^{r-1} \big(F(\,x+\Delta x) - F(x)\big) \big(1 - F(\,x+\Delta x)\big)^{n-r}$$

To get the equation derivative, it will be divided Δx and when Δx limit goes to infinity, $\frac{n!}{(r-1)!(n-r)!}$ will be ignored for simplicity of the next transaction,

$$= \lim_{\Delta x \to \infty} \frac{\big(F(x)\big)^{r-1}\big(F(x+\Delta x) - F(x)\big)\big(1 - F(x+\Delta x)\big)^{n-r}}{\Delta x}$$

Thus,

$$f_{(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!}\big(F(x)\big)^{r-1} f(x)\big(1 - F(x)\big)^{n-r}, \qquad (3.1.1.5)$$

$$f_{(r)}(x) = \frac{1}{\beta(r, n-r+1)}\big(F(x)\big)^{r-1} f(x)\big(1 - F(x)\big)^{n-r}. \qquad (3.1.1.6)$$

3.2 Distribution of Bivariate Order Statistics

Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be two independent identically distributed discrete random variables. $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be a bivariate sample and joint distribution is $F(x, y)$ *with* $F_X(x)$ *and* $F_Y(y)$ margins. $X_{(1)} \leq X_{(2)} \leq, \ldots, \leq X_{(n)}$ be order statistics of $X_1, X_2, \ldots, X_n$ and $Y_{(1)} \leq Y_{(2)} \leq, \ldots, \leq Y_{(n)}$ be order statistics of $Y_1, Y_2, \ldots, Y_n$. In addition, representation of $\big(X_{(r)}, Y_{(s)}\big)$ *r* th and *s* th of bivariate order statistics of X and Y and alternatively can be shown $(X_{r:n}, Y_{s:n})$ where $1 \leq r \leq n \; and \; 1 \leq s \leq n$.

Cumulative distribution function of *r* th and *s* th order statistics are written as binomial distribution of *X* and *Y* respectively,

$$F_{(r)}(x) = P\{X_{(r)} \leq x\} = \sum_{i=r}^{n} C_n^i (F_X(x))^i (1 - F_X(x))^{n-i} = \sum_{i=r}^{n} P\{X = i\} \qquad (3.2.1)$$

$$F_{(s)}(y) = P\{Y_{(s)} \leq y\} = \sum_{j=s}^{n} C_n^j \big(F_Y(y)\big)^j \big(1 - F_Y(y)\big)^{n-j} = \sum_{j=s}^{n} P\{Y = j\} \qquad (3.2.2)$$

In this case, *r* th and *s* th order of bivariate order statistics will be

$$F_{(r),(s)}(x,y) = P\{X_{(r)} \le x\,, Y_{(s)} \le y\} = \sum_{j=s}^{n}\sum_{i=r}^{n} P\{\xi_1 = i\,, \xi_2 = j\} \tag{3.2.3}$$

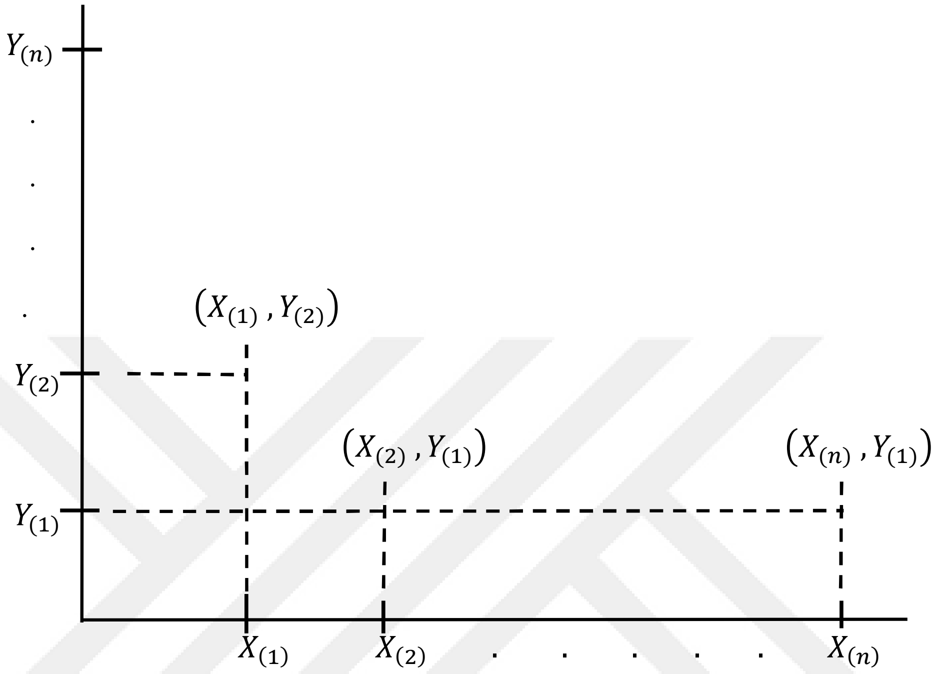

Figure 3.2: Bivariate Order Statistics on Cartesian Coordinate System

We discussed the binomial distribution and the bivariate binomial in detail and $P(i,j) = P\{\xi_1 = i\,, \xi_2 = j\}$ was derived chapter 2 as,

$$= \sum_{h=max(i+j-n,\ 0)}^{min(i,j)} \frac{n!}{h!\,(i-h)!\,(j-h)!\,(n-i-j+h)!}\,\pi_{11}{}^{h}\,\pi_{12}{}^{i-h}\,\pi_{21}{}^{j-h}\,\pi_{22}{}^{n-i-j+h}$$

Thus, $P\{X_{r:n} \le x\,, Y_{s:n} \le y\}$ will be,

$$\sum_{j=s}^{n}\sum_{i=r}^{n}\sum_{h=a}^{b} \frac{n!}{h!\,(i-h)!\,(j-h)!\,(n-i-j+h)!}\,\pi_{11}^{h}\,\pi_{12}{}^{i-h}\,\pi_{21}{}^{j-h}\,\pi_{22}{}^{n-i-j+h} \tag{3.2.4}$$

where

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

and

$$i = 0, \ldots, n; \; j = 0, \ldots, n$$

The two reference books, A First Course in Order Statistics (Arnold, Balakrishnan, and Nagaraja, 2008) and Order Statistics (David, and Nagaraja, 2004) can examine these issues further for those who are interested.

CHAPTER 4: TRIVARIATE BINOMIAL DISTRIBUTION AND TRIVARIATE ORDER STATISTICS

Chapter 4 is the main topic of the thesis and consists of five sub-sections. Based on the bivariate binomial distribution, trivariate binomial distribution will be created in the first subsection. In the following, modification of the trivariate binomial distribution will be discussed. Subsequently, trivariate order statistics will be derived by trivariate binomial distribution. In the fourth subsection, Gumbel copula which is a member of the Archimedean copula family, will be applied to the trivariate order statistics distribution. Finally, the subject will be examined and expanded in detail by giving numerical and graphical examples in the fifth subsection.

4.1 *Trivariate Binomial Distribution*

Let *X*, *Y*, *Z* be binomial random variables and each random variable has two possible outcomes as beloved,

$$X(\omega) = \begin{cases} 1, & if \;\; \omega \in A \\ 0, & if \;\; \omega \in A^C \end{cases}$$

$$Y(\omega) = \begin{cases} 1, & if \;\; \omega \in B \\ 0, & if \;\; \omega \in B^C \end{cases}$$

$$Z(\omega) = \begin{cases} 1, & if \;\; \omega \in C \\ 0, & if \;\; \omega \in C^C \end{cases}$$

$\{\Omega\,,\mathcal{F}\,,P\}$ be probability space and $X(\omega)\,,Y(\omega)\,,Z(\omega)$ defined discrete random variables in this space where $\omega \in \Omega$ and $A\,,B,C \;\in \mathcal{F}$. In this case, eight situations can be occurred, ABC, AB^CC, A^CBC, A^CB^CC, ABC^C, AB^CC^C, A^CBC^C, $A^CB^CC^C$ with probabilities,

$$P\{X = 1\,, Y = 1\,, Z = 1\} = P(ABC) = \pi_{111}, \tag{4.1.1}$$
$$P\{X = 1\,, Y = 0\,, Z = 1\} = P(AB^CC) = \pi_{121},$$

$$P\{X = 0, Y = 1, Z = 1\} = P(A^C BC) = \pi_{211},$$
$$P\{X = 0, Y = 0, Z = 1\} = P(A^C B^C C) = \pi_{221},$$
$$P\{X = 1, Y = 1, Z = 0\} = P(ABC^C) = \pi_{112},$$
$$P\{X = 1, Y = 0, Z = 0\} = P(AB^C C^C) = \pi_{122},$$
$$P\{X = 0, Y = 1, Z = 0\} = P(A^C BC^C) = \pi_{212},$$
$$P\{X = 0, Y = 0, Z = 0\} = P(A^C B^C C^C) = \pi_{222}.$$

$$\sum \pi_{ijk} = 1, \qquad where\ i,j,k = 1,2.$$

Thus, marginals of A, B and C will be,

$$P\{X = 1\} = P(A) = \pi_{111} + \pi_{121} + \pi_{112} + \pi_{122} \tag{4.1.2}$$
$$P\{Y = 1\} = P(B) = \pi_{111} + \pi_{211} + \pi_{112} + \pi_{212} \tag{4.1.3}$$
$$P\{Z = 1\} = P(C) = \pi_{111} + \pi_{121} + \pi_{211} + \pi_{221} \tag{4.1.4}$$

ξ_1, ξ_2, ξ_3 and ξ_{123} defined number of occurrences of *n* experiment as follows,

$$\xi_1 = \{number\ of\ occurences\ A\ in\ n\ experiment\},$$

$$\xi_2 = \{number\ of\ occurences\ B\ in\ n\ experiment\},$$

$$\xi_3 = \{number\ of\ occurences\ C\ in\ n\ experiment\},$$

$$\xi_{123} = \{number\ of\ occurences\ ABC\ in\ n\ experiment\}.$$

and

$$\xi_1 = \sum_{i=1}^{n} X_i, \xi_2 = \sum_{j=1}^{n} Y_j \text{ and } \xi_3 = \sum_{k=1}^{n} Z_k \quad i,j,k = 1,2,3 \ldots n.$$

Let $\xi_1 = i, \xi_2 = j, \xi_3 = k$ be number of occurrences in *n* trials and the marginals of *X*, *Y* and *Z* binomial distributions will be

$$P\{\xi_1 = i\} = \binom{n}{i} (\pi_{111} + \pi_{121} + \pi_{112} + \pi_{122})^i \tag{4.1.5}$$

$$\times \left(1-(\pi_{111}+\pi_{121}+\pi_{112}+\pi_{122})\right)^{n-i}$$

$$P\{\xi_2 = j\} = \binom{n}{j}(\pi_{111}+\pi_{211}+\pi_{112}+\pi_{212})^{j} \tag{4.1.6}$$
$$\times \left(1-(\pi_{111}+\pi_{211}+\pi_{112}+\pi_{212})\right)^{n-j}$$

$$P\{\xi_3 = k\} = \binom{n}{k}(\pi_{111}+\pi_{121}+\pi_{211}+\pi_{221})^{k} \tag{4.1.7}$$
$$\times \left(1-(\pi_{111}+\pi_{121}+\pi_{211}+\pi_{221})\right)^{n-k}$$

When trivariate binomial distribution are shown in the fourfold scheme, *B* and *C* are taken together as one variable to illustrate in one dimension. In this case, there will be four possible events ABC , $A^C BC$, $A(BC)^C$, $A^C(BC)^C$ with $P(ABC) = p_1$, $P(A^C BC) = p_2$, $P(A(BC)^C) = p_3$, $P(A^C(BC)^C) = p_4$ probabilities.

If De Morgan's law is applied to the $A(BC)^C$ and $A^C(BC)^C$,

$$A(BC)^C = A(B \cap C)^C = A(B^C \cup C^C) = AB^C \cup AC^C \tag{4.1.8}$$

$$A^C(BC)^C = A^C(B \cap C)^C = A^C(B^C \cup C^C) = A^C B^C \cup A^C C^C \tag{4.1.9}$$

where

$$P(AB^C) = P(AB^C C) + P(AB^C C^C),$$

$$P(AC^C) = P(ABC^C) + P(AB^C C^C),$$

$$P(A^C B^C) = P(A^C B^C C) + P(A^C B^C C^C),$$

$$P(A^C C^C) = P(A^C BC^C) + P(A^C B^C C^C).$$

p_1 , p_2 , p_3 and p_4 defined as follows,

$$p_1 = P(ABC) \Rightarrow \pi_{111} \tag{4.1.10}$$

$$p_2 = P(A^C BC) \Rightarrow \pi_{211} \tag{4.1.11}$$

$$\begin{aligned} p_3 &= P(A(BC)^C) = P(AB^C \cup AC^C) \\ &= P(AB^C) + P(AC^C) - P(AB^C C^C) \\ &= P(AB^C C) + P(ABC^C) + P(AB^C C^C) \\ &\Rightarrow \pi_{121} + \pi_{112} + \pi_{122} \end{aligned} \tag{4.1.12}$$

$$\begin{aligned} p_4 &= P(A^C(BC)^C) = P(A^C B^C \cup A^C C^C) \\ &= P(A^C B^C) + P(A^C C^C) - P(A^C B^C C^C) \\ &= P(A^C B^C C) + P(A^C BC^C) + P(A^C B^C C^C) \\ &\Rightarrow \pi_{221} + \pi_{212} + \pi_{222} \end{aligned} \tag{4.1.13}$$

To show that the summation of the four probabilities equal to one,

$$P(ABC) + P(A^C BC) + P(A(BC)^C) + P(A^C(BC)^C)$$

$$= P(ABC) + P(A^C BC) + P(AB^C \cup AC^C) + P(A^C B^C \cup A^C C^C)$$

$$= P(ABC) + P(A^C BC) + P(AB^C) + P(AC^C) - P(AB^C C^C)$$

$$+P(A^C B^C) + P(A^C C^C) - P(A^C B^C C^C)$$

$$= P(ABC) + P(A^C BC) + P(AB^C C) + P(A^C B^C C)$$

$$+ P(ABC^C) + P(AB^C C^C) + P(A^C BC^C) + P(A^C B^C C^C) = 1$$

(4.1.14)

Table 4.1: Trivariate Binomial Distribution on Fourfould Scheme

BC | A	A	A^C
BC	ABC p_1 *h times*	A^C BC p_2 *min(j,k)-h times*
$(BC)^C$	$A(BC)^C$ p_3 *i-h times*	$A^C(BC)^C$ p_4 *n-i-min(j,k)+h times*

From n experiments, if *A*, *B* and *C* occurred together *h* times, then they can be figure out $\binom{n}{h}$ ways. On the other hand, *B* and *C* happen together *min(j,k)*. Thus, A^CBC can happened together *min(j,k)-h* times, therefore they can be realized $\binom{n-i}{\min(j,k)-h}$ ways. Since, *A* has a i observation, $A(BC)^C$ can be observed *h-r* times and it can be realized $\binom{n-h}{i-h}$ ways. In addition, $A^C(BC)^C$ can be observed *n-i- min(j,k)+h* times and it can be figure out $\binom{n-i-\,min(j,k)+h}{n-i-\,min(j,k)+h}$ ways. It is obvious that, number of occurrences A ,B ,C are known and the outcomes ABC upper bound must be $min(i, j, k)$ and the lower bound be $max(0\,, i + min(j, k) - n)$ observations. After identifying probabilities, margins, and observations, trivariate binomial distribution will be $P = \{\xi_1 = i, \xi_2 = j,\ \xi_3 = k\}$

$$= \sum_{h=max(0\,,i+min(j,k)-n)}^{min(i,j,k)} \frac{n!}{h!\,(i-h)!\,(min(j,k)-h)!\,(n-i-min(j,k)+h)!}$$

$$\times \; {p_1}^{h} \, {p_2}^{min(j,k)-h} \, {p_3}^{i-h} \, {p_4}^{n-i-min(j,k)+h} \qquad (4.1.15)$$

where,

$$i = 0,\ldots,n \, ; \; j = 0,\ldots,n \, ; \; k = 0,\ldots,n$$

Thus, $min(j,k)$ can take place in three ways, $j < k, k > j \; and \; j = k$. If trivariate binomial distribution is rearranged according to three conditions, it would be,

$$\sum_{h=max\,(0,i+j-n)}^{min\,(i,j)} \frac{n!}{h!\,(i-h)!\,(j-h)!\,(n-i-j+h)!} {p_1}^{h} {p_2}^{j-h} {p_3}^{i-h} {p_4}^{n-i-j+h} \quad ,j < k$$

(4.1.16)

$$\sum_{h=max\,(0,i+k-n)}^{min\,(i,k)} \frac{n!}{h!\,(i-h)!\,(k-h)!\,(n-i-k+h)!} {p_1}^{h} {p_2}^{k-h} {p_3}^{i-h} {p_4}^{n-i-k+h} ,j > k$$

(4.1.17)

$$\sum_{h=max\,(0,i+j-n)}^{min\,(i,j)} \frac{n!}{h!\,(i-h)!\,(j-h)!\,(n-i-j+h)!} {p_1}^{h} {p_2}^{j-h} {p_3}^{i-h} {p_4}^{n-i-j+h} \quad ,j = k$$

(4.1.18)

where $i = 0,\ldots,n \, ; \; j = 0,\ldots,n \, ; \; k = 0,\ldots,n$

4.2 Modification of Trivariate Binomial Distribution

Let *X*, *Y*, *Z* be binomial random variables and each random variable has three possible outcomes A_1, A_2, A_3 for *X*, B_1, B_2, B_3 for *Y* and C_1, C_2, C_3 for *C*. Subsequently, there were twenty-seven possible results which are,

$A_1B_1C_1, A_2B_1C_1, A_3B_1C_1, A_1B_2C_1, A_2B_2C_1, A_3B_2C_1, A_1B_3C_1, A_2B_3C_1, A_3B_3C_1,$

$A_1B_1C_2, A_2B_1C_2, A_3B_1C_2, A_1B_2C_2, A_2B_2C_2, A_3B_2C_2, A_1B_3C_2, A_2B_3C_2, A_3B_3C_2,$

$A_1B_1C_3$, $A_2B_1C_3$, $A_3B_1C_3$, $A_1B_2C_3$, $A_2B_2C_3$, $A_3B_2C_3$, $A_1B_3C_3$, $A_2B_3C_3$, $A_3B_3C_3$.

Moreover, probabilities of each possible results are $\pi_{111}, \pi_{211}, \pi_{311}, \pi_{121}, \pi_{221}, \pi_{321}, \pi_{131}, \pi_{231}, \pi_{331}, \pi_{112}, \pi_{212}, \pi_{312}, \pi_{122}, \pi_{222}, \pi_{322}, \pi_{132}, \pi_{232}, \pi_{332}, \pi_{113}, \pi_{213}, \pi_{313}, \pi_{123}, \pi_{223}, \pi_{323}, \pi_{133}, \pi_{233}, \pi_{333}$ and,

$$\sum \pi_{ijk} = 1, \qquad where\ i,j,k = 1,2,3.$$

For Fourfold scheme, if B and C are taken together as in the subsection 4.1, there are six possible results, $A_1B_1C_1, A_2B_1C_1, A_3B_1C_1, A_1(B_1C_1)^C, A_2(B_1C_1)^C, A_2(B_1C_1)^C$. The probabilities of each occurrence are identified by applying De Morgan's law,

$$p_1 = P(A_1B_1C_1) \Rightarrow \pi_{111} \tag{4.2.1}$$

$$p_2 = P(A_2B_1C_1) \Rightarrow \pi_{211} \tag{4.2.2}$$

$$p_3 = P(A_3B_1C_1) \Rightarrow \pi_{311} \tag{4.2.3}$$

$$p_4 = A_1(B_1C_1)^C = A_1(B_1 \cap C_1)^C = A_1\left({B_1}^C \cup {C_1}^C\right) \tag{4.2.4}$$

$$= P(A_1{B_1}^C) \cup P(A_1{C_1}^C)$$

$$= P\left(A_1{B_1}^C\right) + P\left(A_1{C_1}^C\right) - P(A_1B_2C_2) - P(A_1B_2C_3) - P(A_1B_3C_2) - P(A_1B_3C_3)$$

$$= P(A_1B_2C_1) + P(A_1B_3C_1) + P(A_1B_1C_2) + P(A_1B_1C_3)$$

$$+P(A_1B_2C_2) + P(A_1B_2C_3) + P(A_1B_3C_2) + P(A_1B_3C_3)$$

$$p_4 \Rightarrow \pi_{121}, \pi_{131}, \pi_{112}, \pi_{113}, \pi_{122}, \pi_{123}, \pi_{132}, \pi_{133}$$

where

$$P\left(A_1{B_1}^C\right) = P(A_1B_2C_1) + P(A_1B_3C_1) + P(A_1B_2C_2)$$

$$+P(A_1B_2C_3) + P(A_1B_3C_2) + P(A_1B_3C_3)$$

and

$$P(A_1{C_1}^C) = P(A_1B_1C_2) + P(A_1B_1C_3) + P(A_1B_2C_2)$$

$$+P(A_1B_2C_3) + P(A_1B_3C_2) + P(A_1B_3C_3)$$

$$p_5 = A_2(B_1C_1)^C = A_2(B_1 \cap C_1)^C = A_2\left({B_1}^C \cup {C_1}^C\right) \tag{4.2.5}$$

$$= P(A_2{B_1}^C) \cup P(A_2{C_1}^C)$$

$$= P\left(A_2{B_1}^C\right) + P\left(A_2{C_1}^C\right) - P(A_2B_2C_2) - P(A_2B_2C_3) - P(A_2B_3C_2) - P(A_2B_3C_3)$$

$$= P(A_2B_2C_1) + P(A_2B_3C_1) + P(A_2B_1C_2) + P(A_2B_1C_3)$$

$$+P(A_2B_2C_2) + P(A_2B_2C_3) + P(A_2B_3C_2) + P(A_2B_3C_3)$$

$$p_5 \Rightarrow \pi_{221}, \pi_{231}, \pi_{212}, \pi_{213}, \pi_{222}, \pi_{223}, \pi_{232}, \pi_{233}$$

where

$$P\left(A_2{B_1}^C\right) = P(A_2B_2C_1) + P(A_2B_3C_1) + P(A_2B_2C_2)$$

$$+P(A_2B_2C_3) + P(A_2B_3C_2) + P(A_2B_3C_3)$$

and

$$P(A_2{C_1}^C) = P(A_2B_1C_2) + P(A_2B_1C_3) + P(A_2B_2C_2)$$

$$+P(A_2B_2C_3) + P(A_2B_3C_2) + P(A_2B_3C_3)$$

$$p_6 = A_3(B_1C_1)^C = A_3(B_1 \cap C_1)^C = A_3\left({B_1}^C \cup {C_1}^C\right) \tag{4.2.6}$$

$$= P(A_3{B_1}^C) \cup P(A_3{C_1}^C)$$

$$= P\left(A_3{B_1}^C\right) + P\left(A_3{C_1}^C\right) - P(A_3B_2C_2) - P(A_3B_2C_3) - P(A_3B_3C_2) - P(A_3B_3C_3)$$

$$= P(A_3B_2C_1) + P(A_3B_3C_1) + P(A_3B_1C_2) + P(A_3B_1C_3)$$

$$+P(A_3B_2C_2) + P(A_3B_2C_3) + P(A_3B_3C_2) + P(A_3B_3C_3)$$

$$p_6 \Rightarrow \pi_{321}, \pi_{331}, \pi_{312}, \pi_{313}, \pi_{322}, \pi_{323}, \pi_{332}, \pi_{333}$$

where

$$P(A_3B_1{}^C) = P(A_3B_2C_1) + P(A_3B_3C_1) + P(A_3B_2C_2)$$

$$+P(A_3B_2C_3) + P(A_3B_3C_2) + P(A_3B_3C_3)$$

and

$$P(A_3C_1{}^C) = P(A_3B_1C_2) + P(A_3B_1C_3) + P(A_3B_2C_2)$$

$$+P(A_3B_2C_3) + P(A_3B_3C_2) + P(A_3B_3C_3)$$

and

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1$$

Table 4.2: Modification of Trivariate Binomial Distribution

$\mathrm{B_1C_1 \vert A}$	A_1	A_2	A_3
B_1C_1	$A_1B_1C_1$ p_1 *h times*	$A_2B_1C_1$ p_2 *r times*	$A_3B_1C_1$ p_3 *min(j,k)-h-r times*
$(B_1C_1)^C$	$A_1(B_1C_1)^C$ p_4 *i-h times*	$A_2(B_1C_1)^C$ p_5	$A_3(B_1C_1)^C$ p_6

Let ξ_1 defined the number of times *n* experiment observed A_1, and ξ_2, ξ_3, ξ_{12} are number of occurrences in *n* trial of B_1, C_1 *and* $A_2B_1C_1$, respectively. Assume that $\xi_1 = i, \xi_2 = j, \xi_3 = k, \xi_{12} = r$. From *n* experiments, A_2 and B_1C_1 occurred together *r* times and they can be figure out $\binom{n}{r}$ ways. A_1 and B_1C_1 observed together *h* times, therefore they can be realized $\binom{n-r}{h}$ ways. Since, B_1C_1 has a total of *min(j,k)* observation, A_3 and B_1C_1 occurred together *min(j,k)-h-r* times and they can be realized $\binom{n-r-h}{min\,(j,k)-r-h}$ ways. Moreover, A_2 and B_1 occurred together m times, therefore it can be realized $\binom{n-i}{m}$ ways. A_1 has a i observations, therefore it can be appeared together $(B_1C_1)^C$ with i-h times, and they can be figure out $\binom{n-min\,(j,k)}{i-h}$ ways. It is obvious that, when $A_1B_1C_1$, $A_2B_1C_1$, A_1 and B_1C_1 occurrences known, the outcomes of $A_1B_1C_1$ must be appeared upper bound $min(i, j-r, k-r)$ and the lower bound $max(i + \min\,(j,k) - n,\ \ 0)$ observations.

Under these conditions, $P(i,j,r,m)\ = P\,\{\xi_1\ = i\,, \xi_2\ = j\,, \xi_{12}\ = r\,\}$ will be

$$= \sum_{h=b}^{a} C(i,j,k,r,n)\ p_1{}^{h}\, p_2{}^{r}\, p_3{}^{\min(j,k)-r-h}\, p_4{}^{i-h}(p_5+p_6)^{n-i-\mathrm{m}\quad (j,k)+h}$$

(4.2.7)

where

$$i = 0,\ldots,n\,;\ j\ = r,\ldots,n\,; k\ = r,\ldots,n\,;\ r = 0,\ldots,n$$

and

$$a\ =\ max(0\,, i + min(j,k) - n)\,, b\ =\ \min(i, j-r, k-r)$$

and

$$C(i,j,k,r,n)\ =\ \binom{n}{r}\binom{n-r}{h}\binom{n-r-h}{\min\,(j,k)-r-h}\binom{n-\min\,(j,k)}{i-h}$$

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

4.3 Trivariate Order Statistics

Let the margins A, B, C be defined less than i, j and k, respectively with notation $A = \{X_i \le\ x\}\,, B = \{Y_j \le\ y\}\,, C = \{Z_k \le\ z\}$. Distribution of each probability will be,

$$P(ABC) = P\{X_i \le\ x\,, Y_j \le\ y\ ,\ Z_k \le\ z\}, \qquad (4.3.1)$$

$$P(AB^CC) = P\{X_i \le\ x\,, Y_j \ge\ y\ ,\ Z_k \le\ z\},$$

$$P(A^C BC) = P\{X_i \geq x, Y_j \leq y, Z_k \leq z\},$$
$$P(A^C B^C C) = P\{X_i \geq x, Y_i \geq y, Z_k \leq z\},$$
$$P(ABC^C) = P\{X_i \leq x, Y_j \leq y, Z_k \geq z\},$$
$$P(A^C BC^C) = P\{X_i \geq x, Y_j \leq y, Z_k \geq z\},$$
$$P(AB^C C^C) = P\{X_i \leq x, Y_j \geq y, Z_k \geq z\},$$
$$P(A^C B^C C^C) = P\{X_i \geq x, Y_j \geq y, Z_k \geq z\}.$$

Under these conditions, if trivariate binomial distribution is applied to trivariate order statistics, each probabilities of this distribution as follows,

$$P(ABC) = P\{X_i \leq x, Y_j \leq y, Z_k \leq z\} \Longrightarrow p_1$$
$$= F_{X,Y,Z}(x, y, z) \tag{4.3.2}$$

$$P(A^C BC) = P\{X_i \geq x, Y_j \leq y, Z_k \leq z\} \Longrightarrow p_2$$
$$= F_{Y,Z}(y, z) - F_{X,Y,Z}(x, y, z) \tag{4.3.3}$$

$$P(A(BC)^C) = P(AB^C C) + P(ABC^C) + P(AB^C C^C)$$
$$= P\{X_i \leq x, Y_j \geq y, Z_k \leq z\} + P\{X_i \leq x, Y_j \leq y, Z_k \geq z\}$$
$$+ P\{X_i \leq x, Y_j \geq y, Z_k \geq z\} \Longrightarrow p_3$$
$$= F_X(x) - F_{X,Y,Z}(x, y, z) \tag{4.3.4}$$

$$P(A^C(BC)^C) = P\{X_i \geq x, Y_j \geq y, Z_k \leq z\} + P\{X_i \geq x, Y_j \leq y, Z_k \geq z\}$$
$$+ P\{X_i \geq x, Y_j \geq y, Z_k \geq z\} \Longrightarrow p_4$$
$$= 1 - F_X(x) - F_{Y,Z}(y, z) + F_{X,Y,Z}(x, y, z) \tag{4.3.5}$$

Once the distribution of each probability is defined as above, n samples of r th, s th and t th of trivariate order statistics equal to

$$P = \{X_{r:n} \leq x,\ Y_{s:n} \leq y,\ Z_{t:n} \leq z\}$$

$$= \sum_{i=r}^{n} \sum_{j=s}^{n} \sum_{k=t}^{n} P = \{\xi_1 = i, \xi_2 = j,\ \xi_3 = k\}$$

$$1 \leq r \leq n\ , 1 \leq s \leq n \text{ and } 1 \leq t \leq n.$$

When the trivariate binomial distribution is put into the equation,

$$= \sum_{i=r}^{n} \sum_{j=s}^{n} \sum_{k=t}^{n} \sum_{h=a}^{b} \frac{n!}{h!\,(i-h)!\,(min(j,k)-h)!\,(n-i-min(j,k)+h)!}$$

$$\times\ p_1{}^{h}\, p_2{}^{min(j,k)-h}\, p_3{}^{i-h}\, p_4{}^{n-i-min(j,k)+h} \qquad (4.3.6)$$

where

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

and

$$i,j,k = 0,\ldots,n.$$

4.4 Particular Case: Gumbel Copula

Copula is basically a joint distribution of random variables and two-dimensional copula is a function $C: [0,1]^2 \Rightarrow [0,1]$ with following three properties (Nelsen, 1999):

1. For every $u, v \in [0,1]$, $C(u,0) = C(0,v) = 0$
2. For every $u, v \in [0,1]$, $C(u,1) = u$ and $C(1,v) = v$
3. For every $(u_1, u_2), (v_1 v_2) \in [0,1] \times [0,1]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(v_1, v_2) - C(v_1, u_2) - C(u_1, v_2) + C(u_1, u_2) \geq 0$$

Gumbel copula is a member of the Archimedean copula family and denoted $C_{\theta}^{GU}(u_1, \ldots\ldots, u_n)$. Furthermore, its distribution as follow,

$$C_{\theta}^{GU}(u_1, \ldots\ldots, u_n) = \exp\left(-[\,(-\ln u_1\,)^{\theta} + \ldots + (-\ln u_n)^{\theta}]^{\frac{1}{\theta}}\right),$$

$$\theta \geq 1\,,\ \psi_{\theta}(u) = (-\ln u)^{\theta}.$$

θ shows the degree of dependency, when $\theta = 0$ means that independence copula and $\theta = \infty$ means comonotonicity copula. If the Gumbel copula is applied to the trivariate order statistics distribution with margins $F_X(x) = u$, $F_Y(y) = v$ and $F_Z(z) = w$, trivariate binomial order statistics will be,

$$= \sum_{i=r}^{n}\sum_{j=s}^{n}\sum_{k=t}^{n}\sum_{h=a}^{b} \frac{n!}{h!\,(i-h)!\,(min(j,k)-h)!\,(n-i-min(j,k)+h)!}$$

$$\times\ {p_1}^{h}\,{p_2}^{min(j,k)-h}\,{p_3}^{i-h}\,{p_4}^{n-i-min(j,k)+h} \tag{4.4.1}$$

where,

$$p_1 = C\big(F_X(x), F_Y(y), F_Z(z)\big) \tag{4.4.2}$$

$$= e^{-[-ln(F_X(x))^{\theta} - l\ \ (F_Y(y))^{\theta} - ln(F_Z(z))^{\theta}]^{\frac{1}{\theta}}}$$

$$p_2 = C\big(F_Y(y), F_Z(z)\big) - C\big(F_X(x), F_Y(y), F_Z(z)\big) \tag{4.4.3}$$

$$= e^{-[-ln(F_Y(y))^{\theta} - ln(F_Z(z))^{\theta}]^{\frac{1}{\theta}}} - e^{-[-ln(F_X(x))^{\theta} - ln(F_Y(y))^{\theta} - l\ \ (F_Z(z))^{\theta}]^{\frac{1}{\theta}}}$$

$$p_3 = F_X(x) - C\big(F_X(x), F_Y(y), F_Z(z)\big) \tag{4.4.4}$$

$$= F_X(x) - e^{-[-ln(F_X(x))^{\theta} - ln(F_Y(y))^{\theta} - ln(F_Z(z))^{\theta}]^{\frac{1}{\theta}}}$$

$$p_4 = 1 - F_X(x) - C\big(F_Y(Y), F_Z(z)\big) + C(F_X(x), F_Y(y), F_Z(z)) \tag{4.4.5}$$

$$= 1 - F_X(x) - e^{-[-1\ \ (F_Y(y))^{\theta} - \ln(F_Z(z))^{\theta}]^{\frac{1}{\theta}}} + e^{-[-ln(F_X(x))^{\theta} - ln(F_Y(y))^{\theta} - ln(F_Z(z))^{\theta}]^{\frac{1}{\theta}}}$$

and

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

and

$$i,j,k = 0,\ldots,n.$$

4.5 *Some Numerical Results and Application*

In the first three examples, given probabilities with different theta values were examined. In the fourth example, probabilities are unknown, however the result obtained by re-sampling method. The distribution of the $\theta = 2,\ \theta = 5\ and\ \theta = 8$ values to the three-dimensional Gumbel copula is as follows,

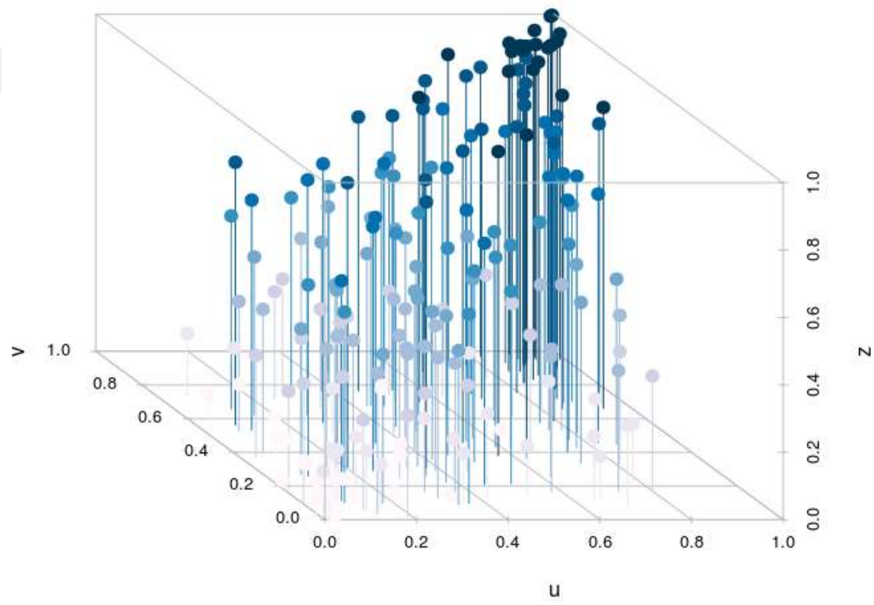

Figure 4.1: Three-dimensional Gumbel copula with θ = 2

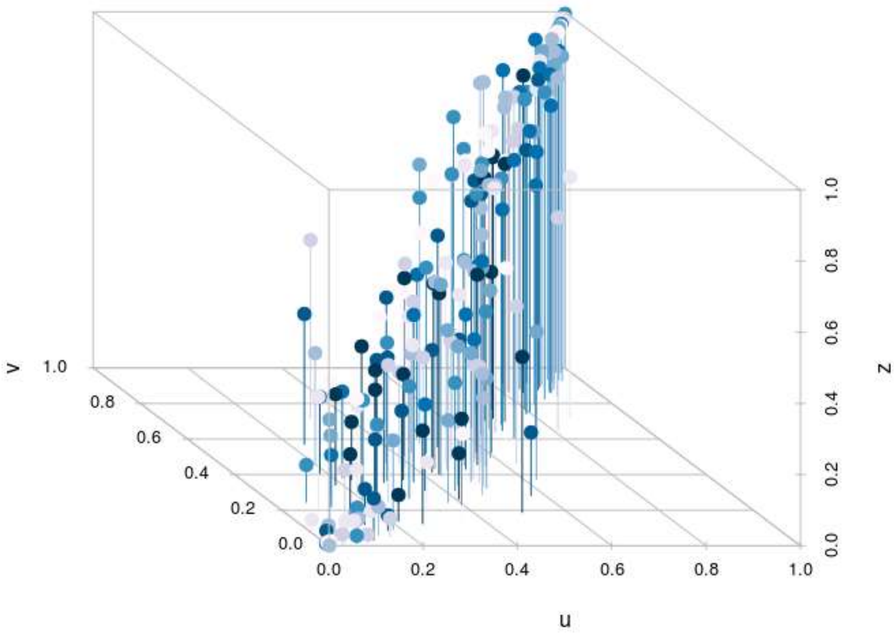

Figure 4.2: Three-dimensional Gumbel copula with $\theta = 5$

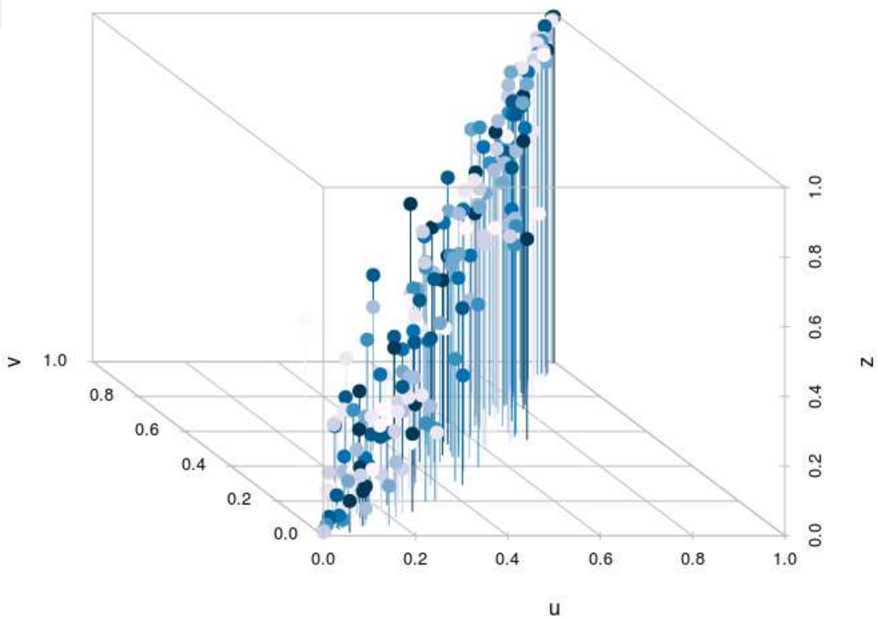

Figure 4.3: Three-dimensional Gumbel copula with $\theta = 8$

4.5.1 Numerical Results I

If probabilities defined as follows,

$$p_1 = 0.08907, \qquad p_3 = 0.51985,$$

$$p_2 = 0.09765, \qquad p_4 = 0.29343.$$

where

$$\pi_{111} = 0.08907, \pi_{121} = 0.14196, \pi_{211} = 0.09765\ , \pi_{221} = 0.03811,$$

$$\pi_{112} = 0.22976\ , \pi_{122} = 0.14813, \pi_{212} = 0.20880, \pi_{222} = 0.04652$$

Table 4.3: Numerical Results I

n	i	j	k	θ	$f(n,i,j,k,\theta)$
8	6	4	4	2	0.024773
				5	5.7069×10^{-5}
				8	3.3536×10^{-7}
5	3	2	3	2	0.008034
				5	1.5723×10^{-5}
				8	9.1954×10^{-8}
6	3	5	4	2	0.000764
				5	1.4624×10^{-7}
				8	6.8413×10^{-11}
4	1	1	3	2	0.143307
				5	0.008435
				8	0.000656

4.5.2 Numerical Results II

If probabilities defined as follows,

$$p_1 = 0.04381, \qquad p_3 = 0.54232,$$

$$p_2 = 0.05423, \qquad p_4 = 0.35964.$$

where

$$\pi_{111} = 0.04381, \pi_{121} = 0.13967, \pi_{211} = 0.05423\ , \pi_{221} = 0.15376,$$

$$\pi_{112} = 0.21777\ , \pi_{122} = 0.18488, \pi_{212} = 0.11906, \pi_{222} = 0.08682$$

Table 4.4 Numerical Results II

n	i	j	k	θ	$f(n,i,j,k,\theta)$
8	6	4	4	2	0.018188
				5	7.3979×10^{-5}
				8	9.9162×10^{-7}
5	3	2	3	2	0.007512
				5	2.2637×10^{-5}
				8	2.9802×10^{-7}
6	3	5	4	2	0.000514
				5	2.9372×10^{-7}
				8	5.861×10^{-10}
4	1	1	3	2	0.155350
				5	0.011693
				8	0.001420

4.5.3 Numerical Results III

If probabilities defined as follows,

$$p_1 = 0.12457, \qquad p_3 = 0.55732,$$

$$p_2 = 0.19479, \qquad p_4 = 0.12332.$$

where

$$\pi_{111} = 0.12457, \pi_{121} = 0.16770, \pi_{211} = 0.19479, \pi_{221} = 0.02589,$$

$$\pi_{112} = 0.20548, \pi_{122} = 0.18414, \pi_{212} = 0.05946, \pi_{222} = 0.03797$$

Table 4.5 Numerical Results III

n	i	j	k	θ	$f(n,i,j,k,\theta)$
8	6	4	4	2	0.032536
				5	8.8258×10^{-5}
				8	1.2108×10^{-6}
5	3	2	3	2	0.006530
				5	1.7097×10^{-5}
				8	2.3986×10^{-7}
6	3	5	4	2	0.000877
				5	4.1377×10^{-7}
				8	8.8079×10^{-10}
4	1	1	3	2	0.084767
				5	0.006521
				8	0.000821

4.5.4 Numerical Results IV

Let n, i, j, k be realized 8,6,4,4, respectively and unlike the other three examples probability values are not known in this case. For the outcome of the distribution, randomly eight probability pairs were selected with the sum of these probabilities were one. By assigning random probability five hundred times, the results were obtained as in Figure 4.4. According to results, $f(n = 8, i = 6, j = 4, k = 4, \theta = 2)$ is occurred between 0.01276 and 0.01858 with a 95% confidence interval.

Table 4.6: Summary Statistics

2.5%	50%	97.5%	Mean
0.01276	0.01555	0.01858	0.01562

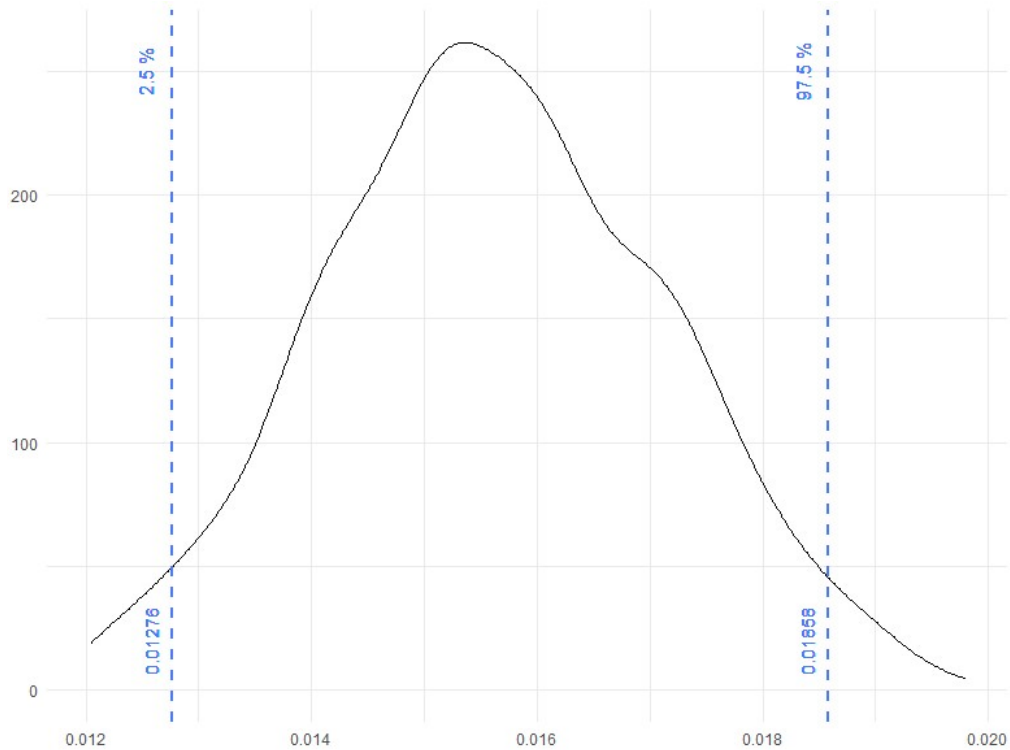

Figure 4.4: Trivariate Order Statistics Density with 500 Random Probability Samples

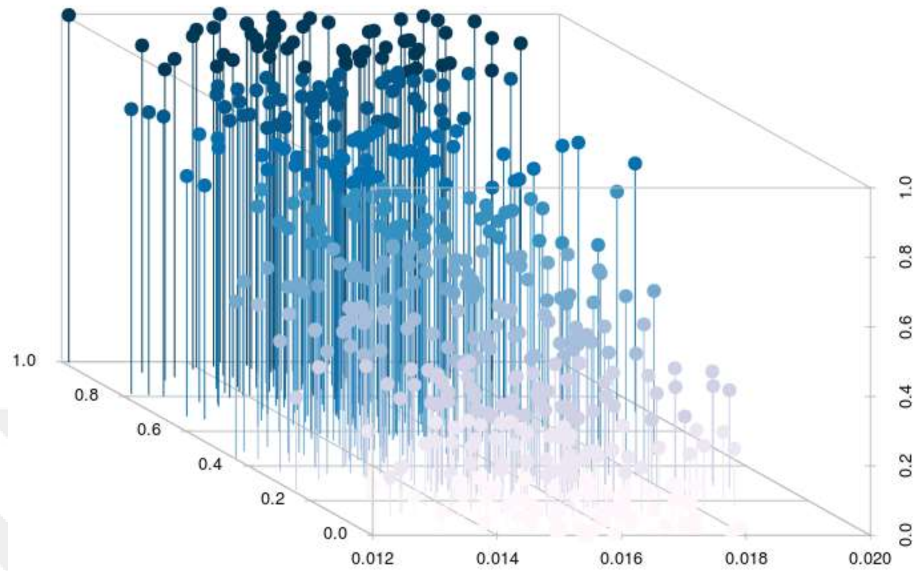

Figure 4.5: Trivariate Order Statistics Cdf with 500 Random Probability Samples

CHAPTER 5: CONCLUSION

Probability generating function, conditional distribution (Papageorgiou, and David, 1994) and linear regression (Chandrasekar, and Balakrishnan, 2002) of trivariate binomial were studied. In this thesis, trivariate binomial distribution is derived by using the fourfold scheme and De Morgan's law. Modifications of trivariate binomial was illustrated and the size of the subject was expanded at the same time. Subsequently, trivariate order statistics were obtained. These new equations can be used in discrete probability models, probability generating functions and many application areas of statistics. Since $(X_1Y_1Z_1), \dots, (X_nY_nZ_n)$ are taken simultaneously as samples of X, Y, Z random variables, it can be integrated to game theory. Furthermore, Gumbel copula has been selected as a special example, because its use extensively in the field of finance to define economic capital adequacy market risk and portfolio analysis. R programming language is used for numerical results and graphical drawings. In cases where the probabilities are not given, it has been shown that this issue can be studied in statistical application areas by re-sampling, confidence intervals and other tools.

REFERENCES

Aitken, A. C. and Gonin, H. T. (1936). *XI.—On Fourfold Sampling with and without Replacement*. Proceedings of the Royal Society of Edinburgh, Vol. 55, pp. 114-125.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (2008). *A first course in order statistics*. 1st edition. Philadelphia: Society for Industrial and Applied Mathematics.

Bairamov, I. and Gultekin, O. E. (2010). *Discrete distributions connected with the bivariate Binomial*. Hacettepe J. Math. Stat, Vol. 39, pp. 109-120.

Barakat, H. M. (1999). *On moments of bivariate order statistics.* Annals of the Institute of Statistical Mathematics, Vol. 51(2), pp. 351-358.

Bairamov, I. and Kemalbay, G. (2013). *Some novel discrete distributions under fourfold sampling schemes and conditional bivariate order statistics*. Journal of Computational and Applied Mathematics, Vol. 248, pp. 1-14.

Bernoulli, J. (1713). *Ars Conjectandi, Opus Posthumum; Accedit Tractatus De Seriebus Infinitis, Et Epistola Gallicè scripta De Ludo Pilae Reticularis*. First edition. Thurnisii.

Biswas, A. and Hwang, J. S. (2002). *A new bivariate binomial distribution*. Statistics & Probability Letters, Vol. 60(2), pp. 231-240.

Boland, P. J., Hollander, M., Joag-Dev, K. and Kochar, S. (1996). *Bivariate dependence properties of order statistics*. Journal of Multivariate Analysis, Vol. 56(1), pp. 75-89.

Chandrasekar, B. and Balakrishnan, A. N. (2002). *Some properties and a characterization of trivariate and multivariate binomial distributions*. Statistics, Vol. 36(3), pp. 211-218.

Crowder, M. and Sweeting T. (1989). *Bayesian inference for a bivariate binomial distribution*. Biometrika, Vol. 76(3), pp. 599–603

David, H. A. and Nagaraja, H. N. (2004). *Order statistics*. 3rd edition. New York: Encyclopedia of Statistical Sciences.

Eryılmaz, S. and Bairamov, I. (2003). *On a new sample rank of an order statistics and its concomitant*. Statistics & Probability Letters, Vol. 63(2), pp. 123-131.

Hamdan, M. A. (1972). *Canonical expansion of the bivariate binomial distribution with unequal marginal indices*. International Statistical Review/Revue Internationale de Statistique, Vol. 40(3), pp. 277-280.

Hamdan, M. A. and Jensen, D. R. (1976), *A bivariate binomial distributions and Some Applications*. Australian Journal of Statistics, Vol. 18, pp. 163-169.

Hamdan, M. A. and Martinson, E. O. (1971), *Maximum Likelihood Estimation in the Bivariate Binomial (0,1) Distribution:Application to 2×2 Tables*. Australian Journal of Statistics, Vol. 13, pp. 154-158.

Kemalbay, G. and Bayramoglu, I. (2015). *Joint distribution of new sample rank of bivariate order statistics*. Journal of Applied Statistics, Vol. 42(10), pp. 2280-2289.

Le, C. T. (1984). *A symmetric bivariate binomial distribution and its application to the analysis of clustered samples in medical research*. Biometrical Journal, Vol. 26(3), pp. 289-294.

Hamdan, M. A. and Nasro M. O. (1986) *Maximum likelihood estimation of the parameters of the bivariate binomial distribution*. Communications in Statistics - Theory and Methods, Vol. 15(3), pp. 747-754.

Marshall, A. W. and Olkin, I. (1985). *A family of bivariate distributions generated by the bivariate Bernoulli distribution*. Journal of the American Statistical Association, Vol. 80(390), pp. 332-338.

Mishra, A. (1996*). A generalised bivariate binomial distribution applicable in four-fold sampling*. Communications in Statistics-Theory and Methods, Vol. 25(8). pp. 1943-1956.

Nelsen, R. B. (2007). *An introduction to copulas*. 2nd edition. Berlin: Springer Science & Business Media.

Oluyede, B. O. (1994). *A family of bivariate binomial distributions generated by extreme Bernoulli distributions*. Communications in Statistics-Theory and Methods, Vol. 23(5), pp. 1531-1547.

Papageorgiou, H. and David, K. M. (1994). *On countable mixtures of bivariate binomial distributions*. Biometrical journal, Vol. 36(5), pp. 581-601.

Papageorgiou, H. and David, K. M. (1995). *The structure of compounded trivariate binomial distributions.* Biometrical journal, Vol. 37(1), pp. 81-95.

Polson, N. and Wasserman, L. (1990). *Prior distributions for the bivariate binomial*. Biometrika, Vol. 77(4), pp. 901-904.

Kocherlakota, S. (1989). *A note on the bivariate binomial distribution*. Statistics & Probability Letters, Vol. 8(1), pp. 21-24.

Teugels, J. L. (1990). *Some representations of the multivariate Bernoulli and binomial distributions.* Journal of Multivariate Analysis, Vol. 32(2), pp. 256-268.

Wilks, S. S. (1942). *Statistical prediction with special reference to the problem of tolerance limits*. The Annals of Mathematical Statistics, Vol. 13(4), pp. 400-409.

Wilks, S. S. (1948). *Order statistics*. Bulletin of the American Mathematical Society, Vol. 54(1), pp. 6-50.

Yule, G. U. (1919). *An introduction to the theory of statistics*. 1st edition. Glasgow: Charles Griffin and Company.

APPENDICES

Code of R for three-dimensional Gumbel copula with $\boldsymbol{\theta = 2}$, $\boldsymbol{\theta = 5}$, $\boldsymbol{\theta = 8}$

```
library(copula)
library(RColorBrewer)
library(scatterplot3d)
set.seed(1994)
gc <- gumbelCopula(2, dim = 3) # theta = 2
U <- rCopula(200, copula = gc)
plotvar <- U[,3]
nclr <- 9
plotclr <- brewer.pal(nclr,"PuBu")
colornum <- cut(rank(plotvar), nclr, labels=FALSE)
colcode <- plotclr[colornum]
plot.angle <- 135
scatterplot3d(U[,1], U[,2], U[,3], type="h", angle=plot.angle, color=colcode,
        pch=20, cex.symbols=2, col.axis="gray", col.grid="gray",
        xlab = "u", ylab = "v", zlab = "z")
gc <- gumbelCopula(5, dim = 3) # theta = 5
U <- rCopula(200, copula = gc)
scatterplot3d(U[,1], U[,2], U[,3], type="h", angle=plot.angle, color=colcode,
```

```
                    pch=20, cex.symbols=2, col.axis="gray", col.grid="gray",
                    xlab = "u", ylab = "v", zlab = "z")
          gc <- gumbelCopula(8, dim = 3) # theta = 8
          U <- rCopula(200, copula = gc)
          scatterplot3d(U[,1], U[,2], U[,3], type="h", angle=plot.angle, color=colcode,
pch=20, cex.symbols=2, col.axis="gray", col.grid="gray", xlab = "u", ylab = "v",
zlab = "z")
```

Code of R for Numerical Example I

```
gumbel.cop <- function(u1,u2,u3,theta){
   c_uvz <- exp(-((-log(u1))^theta + (-log(u2))^theta + (-log(u3))^theta)^(1/theta))
   c_uv <- exp(-((-log(u1))^theta + (-log(u2))^theta)^(1/theta))
   c_uz <- exp(-((-log(u1))^theta + (-log(u3))^theta)^(1/theta))
   c_vz <- exp(-((-log(u2))^theta + (-log(u3))^theta)^(1/theta))
   c_u <- u1
   c_v <- u2
   c_z <- u3
   out <- data.frame(c_uvz,c_uv,c_uz,c_vz,c_u,c_v,c_z)
   return(data.frame(out))
 }
dens_cop <- function(t_cop){
   p_1 <- t_cop$c_uvz
   p_2 <- t_cop$c_vz-t_cop$c_uvz
   p_3 <- t_cop$c_u-t_cop$c_uvz
   p_4 <- 1-t_cop$c_u-t_cop$c_vz+t_cop$c_uvz
   prob_copula <- c(p_1, p_2,p_3,p_4)
```

```
  r_name <- c("p1","p2","p3","p4")
  return(data.frame(prob_copula, row.names = r_name))
}
triv_bin <- function(n,i,j,k,prob1){
  h <- max(0, i + min(j,k) - n)
  density_bin <- (prob1[1,]^h)*(prob1[2,]^(min(j,k)-h))*
   ((prob1[3,])^(i-h))*((prob1[4,])^(n-i-min(j,k)+h))
  fact_bin <- factorial(n)/(factorial(h)*factorial(i-h)*factorial(min(j,k)-h)
*factorial(n-i-min(j,k)+h))
  b <- min(i,j,k)
  sum1 <- (b-h)*(fact_bin*density_bin)
  sum2 <- (n-i)*sum1
  sum3 <- (n-j)*sum2
  sum4 <- (n-k)*sum3
  return(sum4)
 }
p <- c(0.08907, 0.14196,0.09765,0.03811, 0.22976,0.14813,0.20880,0.04652)
prob <- data.frame(p)
u1 <- prob[1,]+prob[2,]+prob[5,]+prob[6,]
u2 <- prob[1,]+prob[3,]+prob[5,]+prob[7,]
u3 <- prob[1,]+prob[2,]+prob[3,]+prob[4,]
for (theta in c(2,5,8)){
 t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
 prob1 <- dens_cop(t_cop)
 print(triv_bin(8,6,4,4,prob1 = prob1))
}
for (theta in c(2,5,8)){
```

```
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(5,3,2,3,prob1 = prob1)) }
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(6,3,5,4,prob1 = prob1))
}
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(4,1,1,3,prob1 = prob1))
}
```

Code of R for Numerical Example II

```
p <- c(0.04381, 0.13967, 0.05423, 0.15376, 0.21777, 0.18488, 0.11906, 0.08682)
prob <- data.frame(p)
u1 <- prob[1,]+prob[2,]+prob[5,]+prob[6,]
u2 <- prob[1,]+prob[3,]+prob[5,]+prob[7,]
u3 <- prob[1,]+prob[2,]+prob[3,]+prob[4,]
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(8,6,4,4,prob1 = prob1))
}
for (theta in c(2,5,8)){
```

```
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(5,3,2,3,prob1 = prob1)) }
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(6,3,5,4,prob1 = prob1))
}
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(4,1,1,3,prob1 = prob1))
}
```

Code of R for Numerical Example III

```
p <- c(0.12457, 0.16770, 0.19479, 0.02589, 0.20548, 0.18414, 0.05946, 0.03797)
prob <- data.frame(p)
u1 <- prob[1,]+prob[2,]+prob[5,]+prob[6,]
u2 <- prob[1,]+prob[3,]+prob[5,]+prob[7,]
u3 <- prob[1,]+prob[2,]+prob[3,]+prob[4,]
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(8,6,4,4,prob1 = prob1))
}
for (theta in c(2,5,8)){
```

```
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(5,3,2,3,prob1 = prob1)) }
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(6,3,5,4,prob1 = prob1))
}
for (theta in c(2,5,8)){
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=theta)
  prob1 <- dens_cop(t_cop)
  print(triv_bin(4,1,1,3,prob1 = prob1))
}
```

Code of R for Numerical Example IV

```
set.seed(11123)
options(digits = 4)
df_total = data.frame()
for (i in 1:500){
  p <- as.numeric(prop.table(table(sample(1:8, size=1000, replace=TRUE))))
  prob <- data.frame(p)
  u1 <- prob[1,]+prob[2,]+prob[5,]+prob[6,]
  u2 <- prob[1,]+prob[3,]+prob[5,]+prob[7,]
  u3 <- prob[1,]+prob[2,]+prob[3,]+prob[4,]
```

```
  t_cop <- gumbel.cop(u1=u1,u2=u2,u3=u3,theta=2)

  prob1 <- dens_cop(t_cop)

  a <- triv_bin(8,6,4,4,prob1 = prob1)

df_total <- rbind(df_total, a)

 }

names(df_total)[1] <- "density"

quantile(df_total$density, probs = c(0.025, 0.50,0.975))

mean(df_total$density)

ggplot(df_total, aes(`density`)) + geom_density()+

geom_vline(xintercept = quantile(df_total$density, probs = c(0.025, 0.975)),

lwd = 0.9, linetype = "dashed", col = "#4d79ff")+

geom_text(aes(x=0.01858, label="97.5 %", y = 250) , colour="#4d79ff", angle=90,
vjust = -1.2, text=element_text(size=14), family = "Helvatica")+

geom_text(aes(x=0.01858 , label="0.01858 ", y = 0) , colour="#4d79ff",angle=90,
vjust = -1,hjust = -0.2, size = 3.7, family = "Helvatica")+

geom_text(aes(x=0.01276, label="2.5 %", y = 250), colour="#4d79ff", angle=90,
vjust = -1.2, text=element_text(size=14), family = "Helvatica")+

geom_text(aes(x=0.01276, label="0.01276 ", y = 0) , colour="#4d79ff",
angle=90,vjust = -1, hjust = -0.2,size = 3.7, family = "Helvatica")+ylab("") + xlab("")
+ theme_minimal()

b <- (df_total$density)

a <- empirical_cdf(b, ubounds=seq(0,1, length.out = 500))

a <- a$UpperBound

plotvar <- seq(0,1,length.out = 500)
```

```
nclr <- 9

plotclr <- brewer.pal(nclr,"PuBu")

colornum <- cut(rank(plotvar), nclr, labels=FALSE)

colcode <- plotclr[colornum]

plot.angle <- 135

s3d <- scatterplot3d(df_total$density, a, plotvar, type="h", angle=plot.angle,
color=colcode, pch=20, cex.symbols=2, col.axis="gray", col.grid="gray")
```