Hypergeometric, or for short GPHD.

The GPHD could not be derived directly from the convolution, because the mixing distribution $\psi(p)$ was defined for positive α_1 and α_2 only. It is worth mentioning that under the circumstances of α_1 and α_2 being non-integers the GPHD can generate negative probabilities and yet have

$$\sum_{x=0}^{S} f(x|S,\alpha_1,\alpha_2) = 1.$$

The limiting form of the BBD and the GPHD is the Binomial distribution if $\alpha_1 \to \infty$ and $\alpha_2 \to \infty$. The BBD plays a major part in modelling of the ascending diagonal arrays. The use of the BBD, GPHD and Binomial distribution for the ascending diagonal arrays makes it possible to have regression curves which are either convex, concave or linear.

The new bivariate and multivariate model enables the creation of a wide range of distributions, just by applying adequate univariate distributions for the third marginal and the ascending diagonal arrays.

1.4 Dirichlet-Multinomial Distribution

The compound multivariate distributions are formed in the same way as the compound univariate probability functions by assigning distributions to some or all of the parameters of the multivariate distribution.

The Dirichlet-Multinomial distribution was obtained from the Multinomial distribution in which the vector of parameters $\underline{p}$ was made to vary in accordance with the Multivariate Beta, which is now called the Dirichlet distribution. Johnson and Kotz (1969) denote the Dirichlet-Multinomial model as follows:

$$\text{Multinomial }(n, p_0, p_1, \ldots, p_k) \underset{p_1 \cdots p_k}{\wedge} \text{ Dirichlet }(\alpha_1, \alpha_2, \ldots, \alpha_k)$$

where (1.4.1)

$$p_0 = 1 - \sum_{i=1}^{k} p_i \ .$$

The probability law of Dirichlet-Multinomial is given by:

$$P(x_1, \ldots, x_k | n, \alpha_0, \alpha_1, \ldots, \alpha_k) = \qquad (1.4.2)$$

$$= \begin{bmatrix} n \\ x_1, \ldots, x_k, \ n - \sum_{i=1}^{k} x_i \end{bmatrix} \frac{\Gamma\left[\sum_{i=0}^{k} \alpha_i\right] \left[\prod_{i=1}^{k} \Gamma\left[x_i + \alpha_i\right]\right] \left[\Gamma\left[n - \sum_{i=1}^{k} x_i + \alpha_0\right]\right]}{\left[\prod_{i=0}^{k} \Gamma\left[\alpha_i\right]\right] \left[\Gamma\left[n + \sum_{i=0}^{k} \alpha_i\right]\right]}$$

where $x_i = 0, 1, 2, \ldots, n$.

This distribution was studied by many authors such as Ishii and Hayakawa (1960), Mosimann (1962), James (1975), Aitchison (1982), Fabius (1973), Firer (1985) and Goodhardt, Ehrenberg and Chatfield (1984).

Mosimann (1962) treated the Dirichlet-Multinomial in a similar manner as Skellam (1948) had developed the Beta-Binomial distribution. In his paper Mosimann discussed some properties of the Dirichlet-Multinomial. He proved the remarkable result that the correlations between any two variables in the

Dirichlet-Multinomial are unchanged from the Multinomial distribution. He obtained also the important result that the correlations in the Dirichlet-Multinomial are equivalent to the correlations in the Dirichlet distribution.

As mentioned by numerous authors these inter-variable correlations are negative and very low. They are quite clearly explained by Mosimann (1962) in his definition of "independence" for proportions, which is *"partial independence except for the constraint"*.

Goodhardt, Ehrenberg and Chatfield (1984) use the Dirichlet-Multinomial as a model of buying behaviour. They developed the buying behaviour model termed as NBD-Dirichlet model by compounding the Dirichlet-Multinomial distribution with the NBD. The assumptions of this model are:

"(a) Purchasing of the product-class takes the form of a Poisson process for each consumer;

(b) The purchasing rates of different consumers follow a Gamma distribution;

(c) Each consumer's choices among the available brands follow a Multinomial distribution; and

(d) These choice probabilities follow a Multivariate Beta or Dirichlet distribution across different consumers."

In discussing Goodhardt, Ehrenberg and Chatfield's paper Kemp (1984) mentioned that the NBD-Dirichlet model can be written shortly as:

$$\left[M(\underline{x}|n,\underline{p}) \underset{\underline{p}}{\wedge} D(\underline{p}|\underline{\alpha})\right] \underset{n}{\wedge} \left[P(n|\lambda) \underset{\lambda}{\wedge} G(\lambda|\gamma,\theta)\right], \qquad (1.4.3)$$

where M, D, P and G denote the Multinomial, Dirichlet, Poisson and Gamma distributions.

Aitchison (1982) stated that *"Dirichlet distribution seldom if ever provides an adequate description of compositional data because of its strong independence structure".*

In this investigation a multivariate model is proposed, which may be characterized by the compounding process:

$$\left[M(\underline{x}|n,\underline{p}) \underset{\underline{p}}{\wedge} D(\underline{p}|\underline{\alpha})\right] \underset{n}{\wedge} F(n|\underline{\eta}), \qquad (1.4.4)$$

where $F(n|\underline{\eta})$ is any family of distributions for n, which is a function of the vector of parameters $\underline{\eta}$.

The above structure creates a wide range of possible correlations and hence it differs from the model presented by Mosimann (1962). However, for a fixed n we end up with Mosimann's (1962) suggestion.

The NBD-Dirichlet model is a particular case of our generalization on the Dirichlet-Multinomial. The possibility to compound on n any discrete family of distributions opens a wide area of applicability of the Dirichlet-Multinomial in compositional data.

1.5 Computing Work

The statistical analysis and the graphical displays have been performed on an IBM-compatible personal computer, using mainly Basic, Lotus 1-2-3 and Statgraphics.

CHAPTER 2

THE BIVARIATE MODEL

This chapter deals with the formulation of a new mathematical model for the bivariate case. The symmetrical case is also considered as a particular situation of the general model.

In order to emphasize the general theory special distributions are applied to the third marginal and to the ascending diagonal arrays. The model enables the use of any distribution and opens the way for the creation of new bivariate probability functions. The latter are not necessarily bivariate versions of a specific univariate distribution. In this way it is possible to build new bivariate discrete distributions which are not Poisson mixtures.

2.1 The General Model

In the classical bivariate cases the marginal distributions of x_1 and x_2 determine the correlations.

In this chapter a new mathematical model is developed. The model is based on a new principle which uses the third marginal distribution and the ascending diagonal arrays.

Let us restate some definitions and theorems needed in the mathematical development.

Definition 2.1

The third marginal is the distribution of a new variable S defined as the summation of the marginal variables x_1 and x_2 ($S = x_1+x_2$). Let us denote the distribution of the third marginal as $f(S)$.

Definition 2.2

The ascending diagonal array probability law is the distribution of x_1 given S, where $x_1 = 0,1,2,\ldots S-1,S$. Let us denote the distribution of the ascending diagonal array as $\phi(x_1|S)$.

The bivariate distributions created by the third marginal and the ascending diagonal arrays can be presented in a bivariate table as shown in Figure 2.1.

Lemma 2.1 (Mood and Graybill, 1974).

Let (x_1,x_2) be a two dimensional random variable and let $g(.)$ be a function of x_2. Then:

$$E[g(x_2)] = E\{E[g(x_2)\,|x_1]\}. \tag{2.1.1}$$

<u>Figure 2.1</u> The bivariate table with x_1 marginal, x_2 marginal and the third marginal

S	f(S)	$x_2 \backslash x_1$	0	1	2	...	n-2	n-1	n	$f(x_2)$
0		0								
1		1								
2		2								
⋮		⋮								
n-2		n-2								
n-1		n-1								
n		n								
		$f(x_1)$								

<u>Lemma 2.2</u> **<u>(Mood and Graybill, 1974)</u>.**

Let (x_1,x_2) be a two dimensional random variable. Then:

$$Var(x_2) = E[Var(x_2|x_1)] + Var[E(x_2|x_1)] \tag{2.1.2}$$

2.1.1 The new model

Theorem 2.1

For any distribution $f(S)$ with finite $\mu_1'(S)$ of the third marginal and for any distribution $\phi(x_1|S)$ with finite $\mu_1'(x_1|S)$ of the ascending diagonal array the first moments around the origin $\mu_1'(x_1)$ and $\mu_1'(x_2)$ for x_1 and x_2 can be expressed as follows:

$$\mu_1'(x_1) = \sum_S \mu_1'(x_1|S)\ f(S) \quad \text{and} \qquad (2.1.3)$$

$$\mu_1'(x_2) = \mu_1'(S) - \sum_S \mu_1'(x_1|S)\ f(S)\ . \qquad (2.1.4)$$

Proof

(a) Using (2.1.1) we can write

$E[E(x_1|S)] = E(x_1)$. Therefore:

$\sum_S \sum_{x_1=0}^{S} [x_1\ \phi(x_1|S)]\ f(S) = \mu_1'(x_1)$. Considering that

$\sum_{x_1=0}^{S} x_1\ \phi(x_1|S)] = \mu_1'(x_1|S)$ we obtain equation (2.1.3).

(b) From the definition of the ascending diagonal array it follows that

$x_2 = S - x_1$ and

$E(x_2) = E(S - x_1)$.

By using (a) we obtain:

$\mu_1'(x_2) = E\{E[(S-x_1)|S]\}$, or

$$\mu_1'(x_2) = \sum_S \left[\sum_{x_1=0}^{S} (S-x_1)\ \phi(x_1|S) \right] f(S)$$

$$= \sum_S \left[\sum_{x_1=0}^{S} S\ \phi(x_1|S) - \sum_{x_1=0}^{S} x_1\ \phi(x_1|S) \right] f(S)$$

$$= \sum_S S\ f(S) - \sum_S \mu_1'(x_1|S)\ f(S)\ .$$

This leads to equation (2.1.4) which can also be written as

$\mu_1'(x_2) = \mu_1'(S) - \mu_1'(x_1)$.

Theorem 2.2

For any distribution $f(S)$ with finite moments $\mu_1'(S)$ and $\mu_2'(S)$ for the third marginal and for any distribution $\phi(x_1|S)$ with finite $\mu_1'(x_1|S)$ and $\mu_2'(x_1|S)$ for the ascending diagonal array the second central moments $\mu_2(x_1)$ and $\mu_2(x_2)$ for x_1 and x_2 can be expressed as follows:

$$\mu_2(x_1) = \sum_S \mu_2'(x_1|S)\ f(S) - \left[\sum_S \mu_1'(x_1|S)\ f(S)\right]^2 \tag{2.1.5}$$

and

$$\mu_2(x_2) = \sum_S \mu_2'(x_1|S)\ f(S) - \left[\sum_S \mu_1'(x_1|S)\ f(S)\right]^2 + \mu_2(S)$$
$$- 2\left[\sum_S S\ \mu_1'(x_1|S)\ f(S) - \mu_1'(S) \sum_S \mu_1'(x_1|S)\ f(S)\right]. \tag{2.1.6}$$

Relation (2.1.6) can also be written as:

$$\mu_2(x_2) = \mu_2(x_1) + \mu_2(S)$$
$$- 2\left[\sum_S S\,\mu_1'(x_1|S)\;f(S) - \mu_1'(S)\sum_S \mu_1'(x_1|S)\;f(S)\right],$$

or (2.1.7)

$$\mu_2(x_2) = \mu_2(x_1) + \mu_2(S) - 2\mathrm{cov}(x_1,S). \qquad (2.1.8)$$

Proof

(a) Using (2.1.2) we can calculate the variance of x_1 as follows:

$$\mathrm{Var}(x_1) = E[\mathrm{Var}(x_1|S)] + \mathrm{Var}[E(x_1|S)]\;, \qquad (2.1.9)$$

$$\mu_2(x_1|S) = \mu_2'(x_1|S) - [\mu_1'(x_1|S)]^2\;,$$

$$E[\mathrm{Var}(x_1|S)] = \sum_S \mu_2'(x_1|S)\;f(S) - \sum_S [\mu_1'(x_1|S)]^2\;f(S) \qquad (2.1.10)$$

$$\mathrm{Var}[E(x_1|S)] = \sum_S [\mu_1'(x_1|S)]^2 f(S) - [\sum_S \mu_1'(x_1|S)f(S)]^2. \qquad (2.1.11)$$

By substituting (2.1.10) and (2.1.11) into (2.1.9) we obtain:

$$\mu_2(x_1) = \sum_S \mu_2'(x_1|S)f(S) - [\sum_S \mu_1'(x_1|S)f(S)]^2.$$

(b) Using (2.1.2) we can calculate the variance of x_2 as follows:

$$\mathrm{Var}(x_2) = E\{\mathrm{Var}[(S-x_1)|S]\} + \mathrm{Var}\{E[(S-x_1)|S]\}. \qquad (2.1.12)$$

From the variance properties one obtains the following:

$$\mathrm{Var}(S-x_1|S) = \mu_2'(x_1|S) - [\mu_1'(x_1|S)]^2,$$

$$E\{Var[(S-x_1)|S]\} = \sum_S \mu_2'(x_1|S)f(S) - \sum_S [\mu_1'(x_1|S)]^2 f(S), \qquad (2.1.13)$$

$E[(S-x_1)|S] = S - \mu_1'(x_1|S)$ and

$$\begin{aligned} Var\{E[(S-x_1)|S]\} = {} & \mu_2'(S) - [\mu_1'(S)]^2 + \sum_S [\mu_1'(x_1|S)]^2 f(S) \\ & - [\sum_S \mu_1'(x_1|S) f(S)]^2 \qquad (2.1.14) \\ & - 2 [\sum_S S \mu_1'(x_1|S) f(S) - \mu_1'(S) \sum_S \mu_1'(x_1|S) f(S)] . \end{aligned}$$

By substituting (2.1.13) and (2.1.14) into (2.1.12) one finds

$$\begin{aligned} \mu_2(x_2) = {} & \sum_S \mu_2(x_2|S) f(S) - [\sum_S \mu_1'(x_1|S) f(S)]^2 + \mu_2'(S) - [\mu_1'(S)]^2 \\ & -2 [\sum_S S \mu_1'(x_1|S) f(S) - \mu_1'(S)\sum_S \mu_1'(x_1|S) f(S)]. \qquad (2.1.15) \end{aligned}$$

This leads to equation (2.1.7).

Theorem 2.3

For any distribution $f(S)$ with finite $\mu_1'(S)$ and $\mu_2'(S)$ for the third marginal and for any distribution $\phi(x_1|S)$ with finite $\mu_1'(x_1|S)$ and $\mu_2'(x_2|S)$ for the ascending diagonal array the $\mu_1'(x_1x_2)$ and $cov(x_1,x_2)$ can be expressed as follows:

$$E(x_1x_2) = \sum_S S \mu_1'(x_1|S) f(S) - \sum_S \mu_2'(x_1|S) f(S) \qquad (2.1.16)$$

and

$$\begin{aligned} cov(x_1,x_2) = {} & \sum_S S \mu_1'(x_1|S) f(S) - \sum_S \mu_2'(x_1|S) f(S) \\ & - [\sum_S \mu_1'(x_1|S) f(S)] [\mu_1'(S) - \sum_S \mu_1'(x_1|S) f(S)] \qquad (2.1.17) \end{aligned}$$

Proof

(a) Using Lemma 2.1 we can write:

$$E(x_1x_2) = E[x_1(S-x_1)] = E\{E[x_1(S-x_1)|S]\}$$

$$= \sum_S [\sum_{x_1} x_1(S-x_1)\phi(x_1|S)] f(S)$$

$$= \sum_S S \mu_1'(x_1|S) f(S) - \sum_S \mu_2'(x_1|S) f(S).$$

(b) From the definition of covariance, Theorem 2.1 and part (a) above it is obvious that the covariance can be written as:

$$\mathrm{cov}(x_1,x_2) = \sum_S S \mu_1'(x_1|S) f(S) - \sum_S \mu_2'(x_1|S) f(S)$$

$$- [\sum_S \mu_1'(x_1|S) f(S)] [\mu_1'(S) - \sum_S \mu_1'(x_1|S) f(S)].$$

From the definition of the correlation coefficient, the covariance and Theorems 2.1 and 2.2 we obtain the formula for $\rho(x_1,x_2)$ which is based on the moments around the origin and the mean of the third marginal and the ascending diagonal array.

2.2 The Symmetrical Model

In this section a particular case of the model presented in Section 2.1 is developed.

Let us assume a bivariate distribution expressed in terms of the third marginal and the ascending diagonal array distributions. In order to obtain a bivariate symmetrical distribution in the bivariate table with respect to the main diagonal (See Figure 2.2), the following condition must be satisfied within each diagonal array:

$$\phi(x_1|S) = \phi(x_2|S) \text{ , where} \tag{2.2.1}$$

$x_2 = S - x_1$.

<u>Figure 2.2</u> The bivariate table with symmetry along the main diagonal

S	f(S)	$x_2 \backslash x_1$	0	1	2	...	n-1	n	$f(x_2)$
0		0	$\phi(0\|0)f(0)$	$\phi(1\|1)f(1)$	$\phi(2\|2)f(2)$	...	$\phi(n-1\|n-1)f(n-1)$	$\phi(n\|n)f(n)$	
1		1	$\phi(0\|1)f(1)$	$\phi(1\|2)f(2)$					
2		2	$\phi(0\|2)f(2)$						
⋮		⋮	⋮						
		n- 1	$\phi(0\|n-1)f(n-1)$						
		n	$\phi(0\|n)f(n)$						
		$f(x_1)$							

Theorem 2.4

For any distribution $f(S)$ with finite $\mu_1'(S)$ and $\mu_2'(S)$ for the third marginal and for any symmetrical distribution $\phi(x_1|S)$ with finite $\mu_1'(x_1|S)$ and $\mu_2'(x_2|S)$ for the ascending diagonal array the following relations for $\mu_1'(x_1)$, $\mu_1'(x_2)$, $\mu_2(x_1)$, $\mu_2(x_2)$ and $\rho(x_1,x_2)$ exist:

$$\mu_1'(x_1) = \mu_1'(x_2) \ , \tag{2.2.2}$$

$$\mu_2(x_1) = \mu_2(x_2) \quad \text{and} \tag{2.2.3}$$

$$\rho(x_1,x_2) = \frac{\sum_S S\,\mu_1'(x_1|S)\,f(S) - \sum_S \mu_2'(x_1|S)\,f(S)}{\sum_S \mu_1'(x_1|S)\,f(S)} - \sum_S \mu_1'(x_1|S)\,f(S). \tag{2.2.4}$$

Proof

(a) Since $\phi(x_1|S)$ is a symmetrical distribution we have

$\phi(x_1|S) = \phi(x_2|S)$, as well as

$\mu_1'(x_1) = \sum_S \mu_1'(x_1|S)\,f(S)$ and

$\mu_1'(x_2) = \sum_S \mu_1'(x_2|S)\,f(S)$.

Hence $\mu_1'(x_1) = \mu_1'(x_2)$.

(b) In a similar way to that of part (a) we conclude that:

$\mu_2(x_1|S) = \sum_S \mu_2'(x_1|S)\,f(S) - [\sum_S \mu_1'(x_1|S)\,f(S)]^2$,

$\mu_2(x_2|S) = \sum_S \mu_2'(x_2|S)\,f(S) - [\sum_S \mu_1'(x_2|S)\,f(S)]^2$ and

$\mu_2(x_1) = \mu_2(x_2)$.

(c) Considering the symmetrical distribution for the ascending diagonal arrays and the results of parts (a) and (b) above one can easily obtain from Theorem 2.3 the expression (2.2.4) for $\rho(x_1,x_2)$.

2.3 Special Distributions for the Ascending Diagonal Arrays

The theory developed in the previous two sections permits the use of any distribution for the third marginal and the ascending diagonal arrays.

From definition 2.2 it is obvious that for the ascending diagonal arrays we can only take distributions which are limited by an upper terminal. For all these arrays one type of distribution can be considered, with:

(a) the same parameters for all the arrays, or

(b) different parameters which are dependent on S.

When working with real data sets we encounter two different types of problems in estimating the parameters.

(a) The estimates for these parameters in each array can be displayed graphically and are dispersed randomly without any relation to the array S. Therefore, a measure can be found to estimate the respective parameters. These parameters are considered constant along the whole bivariate table.

(b) The estimates of the parameters follow a certain pattern as a function of S. In this case we can actually fit a function, so that the parameter varies as a function of S. For this type of problem the model described in Section 2.1 will encounter difficulties. However the model can still be applied for certain functions of the parameters and for certain distributions.

In what follows we concentrate the research on three families of distributions:

(a) Beta-Binomial with constant parameters α_1 and α_2,

(b) Generalized Positive Hypergeometric and

(c) Beta-Binomial with parameters α_1 and α_2 as functions of S.

2.3.1 <u>Beta-Binomial distribution with constant parameters α_1 and α_2</u>

In this chapter we develop a particular model which assumes the Beta-Binomial distribution with the same parameters α_1 and α_2 for all the ascending diagonal arrays.

The Beta-Binomial was derived by compounding the Binomial distribution with parameters S and θ with the continuous Beta distribution with parameters α_1 and α_2. In this case Beta is the distribution of the parameter θ.

The probability law of the Binomial distribution is:

$$P(x_1|S) = \begin{bmatrix} S \\ x_1 \end{bmatrix} \theta^{x_1} (1-\theta)^{S-x_1}, \text{ where} \tag{2.3.1}$$

$x_1 = 0,1,2,\ldots,S,\quad S = 1,2,3,\ldots$ and $0 < \theta < 1$.

The probability law of the Beta distribution is:

$$\varphi(\theta) = \frac{1}{B(\alpha_1,\alpha_2)} \theta^{\alpha_1-1} (1-\theta)^{\alpha_2-1}, \text{ where} \tag{2.3.2}$$

$0 < \theta < 1,\quad \alpha_1 > 0$ and $\alpha_2 > 0$.

From the convolution process we obtain the following probability function:

$$\phi(x_1|S) = \begin{bmatrix} S \\ x_1 \end{bmatrix} \frac{B(x_1+\alpha_1, S-x_1+\alpha_2)}{B(\alpha_1,\alpha_2)}, \text{ where} \tag{2.3.3}$$

$x_1 = 0,1,2,\ldots,S,\quad S = 1,2,3,\ldots,$

$\alpha_1 > 0$ and $\alpha_2 > 0$.

The latter function is called the Beta-Binomial distribution, or for short the BBD.

The first two moments around the origin for the BBD are

$$\mu_1'(x_1|S) = \frac{S\alpha_1}{\alpha_1+\alpha_2} \text{ and} \tag{2.3.4}$$

$$\mu_2'(x_1|S) = \frac{S\alpha_1[S(\alpha_1+1)+\alpha_2]}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}. \tag{2.3.5}$$

The BBD can take a vast variety of shapes depending on the values of its two parameters α_1 and α_2.

The ratio function $R(x|S)$ for the BBD is defined as:

$$R(x_1|S) = \frac{\phi(x_1+1|S)}{\phi(x_1|S)} = \frac{(S-x_1)(x_1+\alpha_1)}{(x_1+1)(S-x_1+\alpha_2-1)} \text{ for} \tag{2.3.6}$$

$x_1 = 0,1,2,\ldots,(S-1)$ and $S = 1,2,\ldots$.

$$P(x_1|S) = \binom{S}{x_1} \theta^{x_1} (1-\theta)^{S-x_1}, \text{ where} \tag{2.3.1}$$

$x_1 = 0,1,2,\ldots,S,\quad S = 1,2,3,\ldots$ and $0 < \theta < 1$.

The probability law of the Beta distribution is:

$$\Psi(\theta) = \frac{1}{B(\alpha_1,\alpha_2)} \theta^{\alpha_1-1} (1-\theta)^{\alpha_2-1}, \text{ where} \tag{2.3.2}$$

$0 < \theta < 1,\quad \alpha_1 > 0$ and $\alpha_2 > 0$.

From the convolution process we obtain the following probability function:

$$\phi(x_1|S) = \binom{S}{x_1} \frac{B(x_1+\alpha_1, S-x_1+\alpha_2)}{B(\alpha_1,\alpha_2)}, \text{ where} \tag{2.3.3}$$

$x_1 = 0,1,2,\ldots,S,\quad S = 1,2,3,\ldots$,

$\alpha_1 > 0$ and $\alpha_2 > 0$.

The latter function is called the Beta-Binomial distribution, or for short the BBD.

The first two moments around the origin for the BBD are

$$\mu_1'(x_1|S) = \frac{S\alpha_1}{\alpha_1+\alpha_2} \text{ and} \tag{2.3.4}$$

$$\mu_2'(x_1|S) = \frac{S\alpha_1[S(\alpha_1+1)+\alpha_2]}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}. \tag{2.3.5}$$

The BBD can take a vast variety of shapes depending on the values of its two parameters α_1 and α_2.

The ratio function $R(x|S)$ for the BBD is defined as:

$$R(x_1|S) = \frac{\phi(x_1+1|S)}{\phi(x_1|S)} = \frac{(S-x_1)(x_1+\alpha_1)}{(x_1+1)(S-x_1+\alpha_2-1)} \text{ for} \tag{2.3.6}$$

$x_1 = 0,1,2,\ldots,(S-1)$ and $S = 1,2,\ldots$.

Most of the literature mentions the necessary, but not the sufficient conditions for the BBD to take various shapes.

In order to know the shape of the BBD one should analyse the ratio function at two points $x_1 = 0$ and $x_1 = S-1$. From this analysis one finds the necessary and sufficient conditions for the BBD to take the following shapes:

(a) U-shaped

$$\alpha_1 > S\alpha_2 - (S-1) \quad \text{and} \quad \alpha_2 > S\alpha_1 - (S-1); \tag{2.3.7}$$

(b) J-shaped

$$\alpha_1 > S\alpha_2 - (S-1) \quad \text{and} \quad \alpha_2 \leq S\alpha_1 - (S-1); \tag{2.3.8}$$

(c) Reverse J-shaped

$$\alpha_1 \leq S\alpha_2 - (S-1) \quad \text{and} \quad \alpha_2 > S\alpha_1 - (S-1); \tag{2.3.9}$$

(d) Bell shaped

$$\alpha_1 \leq S\alpha_2 - (S-1) \quad \text{and} \quad \alpha_2 \leq S\alpha_1 - (S-1); \tag{2.3.10}$$

(e) Rectangular

$$\alpha_1 = 1 \quad \text{and} \quad \alpha_2 = 1; \tag{2.3.11}$$

(f) Mirrored (with respect to x-axis) J-shaped

$$\alpha_1 < S\alpha_2 - (S-1) \quad \text{and} \quad \alpha_2 \geq S\alpha_1 - (S-1) \quad \text{and} \tag{2.3.12}$$

(g) Mirrored (with respect to x-axis) reverse J-shaped

$$\alpha_1 \geq S\alpha_2 - (S-1) \quad \text{and} \quad \alpha_2 < S\alpha_1 - (S-1). \tag{2.3.13}$$

For large values of α_1 and α_2 the BBD can be approximated by the Binomial distribution as follows:

$$\phi(x_1|S) = \frac{S!}{x_1!(S-x_1)!} \frac{[\alpha_1(\alpha_1+1)\ldots(\alpha_1-1+x)][\alpha_2(\alpha_2+1)\ldots(\alpha_2-1+S-x_1)]}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)\ldots(S+\alpha_1+\alpha_2-1)}$$

$$= \frac{S!}{x_1!(S-x_1)!} \frac{\frac{\alpha_1}{\alpha_1+\alpha_2} \cdot \frac{\alpha_1+1}{\alpha_1+\alpha_2} \cdot \ldots \cdot \frac{\alpha_1+x_1-1}{\alpha_1+\alpha_2}}{1 \cdot \frac{\alpha_1+\alpha_2+1}{\alpha_1+\alpha_2} \cdot \ldots \cdot \frac{\alpha_1+\alpha_2+x_1-1}{\alpha_1+\alpha_2}}$$

$$\cdot \frac{\frac{\alpha_2}{\alpha_1+\alpha_2} \cdot \frac{\alpha_2+1}{\alpha_1+\alpha_2} \cdot \ldots \cdot \frac{\alpha_2+S-x_1-1}{\alpha_1+\alpha_2}}{1 \cdot \frac{\alpha_1+\alpha_2+x_1}{\alpha_1+\alpha_2} \cdot \ldots \cdot \frac{\alpha_1+\alpha_2+S-1}{\alpha_1+\alpha_2}} . \qquad (2.3.14)$$

If one takes $\alpha_2 = a\alpha_1$ then:

$$\lim_{\substack{\alpha_1 \to \infty \\ \alpha_2 \to \infty}} \phi(x_1|S) = \begin{bmatrix} S \\ x_1 \end{bmatrix} p^{x_1} (1-p)^{S-x_1} , \text{ where} \qquad (2.3.15)$$

$p = 1/(a+1)$.

This is the Binomial distribution.

Under certain conditions the BBD approaches the Negative Binomial distribution. If we assume that there exist a linear relation between α_2 and S, in other words if

$$\alpha_2 = kS , \text{ then} \qquad (2.3.16)$$

$$\lim_{S \to \infty} \phi(x_1|S) = \frac{(1-\theta)^{\gamma} \, \Gamma(\gamma+x_1)}{\Gamma(\gamma) \, \Gamma(x_1+1)} \theta^{x_1} , \text{ where} \qquad (2.3.17)$$

$x_1 = 0,1,2,\ldots,$ $\theta = 1/(k+1)$, $\gamma = \alpha_1$, $0 < \theta < 1$ and $\gamma > 0$.

The conclusion is that in the limiting process for $S \to \infty$ the BBD, which is a compound Binomial, becomes a compound Poisson, which is the NBD.

2.4 Distributions for the Third Marginal

From the definition 2.1 it is obvious that for the third marginal one can choose any distribution which may or may not be limited by an upper value.

In what follows we investigate the model, where on the third marginal we assume the Generalized Inverse Gaussian-Poisson family. Some well known distributions which are immediately derived from the above family will be used as probability laws for the third marginal, such as the Inverse Gaussian-Poisson and the NBD. The BBD and the GPHD presented previously for the ascending diagonal arrays can also be used for modeling the third marginal.

2.4.1 Generalized Inverse Gaussian-Poisson distribution

For the third marginal we now consider the Generalized Inverse Gaussian-Poisson distribution, or for short GIGP. It is defined as

$$f(S) = \frac{(\sqrt{1-\theta}\,)^{\gamma}}{K_{\gamma}(\alpha\sqrt{1-\theta}\,)} \, \frac{(\alpha\theta/2)^{S}}{S!} \, K_{S+\gamma}(\alpha) \ , \quad \text{where} \qquad (2.4.1)$$

$S = 0,1,2,3,\ldots$, $-\infty < \gamma < \infty$, $0 \le \theta \le 1$, $\alpha \ge 0$, and $K_{\gamma}(z)$ is the modified Bessel function of the second kind of order γ and argument z. This is a discrete distribution obtained through the convolution process of the Poisson distribution given by

$$p(x) = e^{-x}\frac{\lambda^{x}}{x!} \ , \quad \text{where } x = 0,1,2,\ldots \text{ and for } \lambda \ge 0 \ , \qquad (2.4.2)$$

2.3.2 Generalized Positive Hypergeometric distribution

For negative values of the parameters α_1 and α_2 we obtain the Generalized Positive Hypergeometric distribution.

By using the equality

$$\begin{bmatrix} x+k-1 \\ x \end{bmatrix} = (-1)^x \begin{bmatrix} -k \\ x \end{bmatrix} \tag{2.3.18}$$

equation (2.3.1) becomes:

$$\phi(x_1|S) = \frac{\begin{bmatrix} -\alpha_1 \\ x_1 \end{bmatrix} \begin{bmatrix} -\alpha_2 \\ S-x_1 \end{bmatrix}}{\begin{bmatrix} -\alpha_1-\alpha_2 \\ S \end{bmatrix}}, \quad \text{where} \tag{2.3.19}$$

$x_1 = 0,1,\ldots,S$, $S = 1,2,\ldots$, $\alpha_1 > 0$ and $\alpha_2 > 0$.

If α_1 and α_2 are taken as negative the letter expression gives:

$$\phi(x_1|S) = \frac{\begin{bmatrix} \alpha_1 \\ x_1 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ S-x_1 \end{bmatrix}}{\begin{bmatrix} \alpha_1+\alpha_2 \\ S \end{bmatrix}}, \quad \text{where} \tag{2.3.20}$$

$x_1 = 0,1,\ldots,S$ and $S = 1,2,\ldots$.

This function is called the Generalized Positive Hypergeometric distribution, or for short GPHD.

For integer values of α_1 and α_2 equation (2.3.20) becomes the Hypergeometric distribution.

Under certain circumstances, for non-integers α_1 and α_2, GPHD can generate negative probabilities and simultaneously:

$$\sum_{x_1}^{\infty} \phi(x_1|S) = 1 \ .$$

The GPHD could not be derived directly from the convolution because the mixing distribution $\Psi(\theta)$ was defined only for positive values of α_1 and α_2 .

In the limiting process the GPHD becomes the Binomial distribution. The proof is identical to the one performed above for the BBD, the only difference being the signs of α_1 and α_2 .

2.3.3 <u>Beta-Binomial distribution with α_1 and α_2 as functions of S</u>

Let us assume a special case where on the ascending diagonal array the variable is distributed according to the BBD, such that the parameters α_1 and α_2 are functions of S.

We assume that the following relation exists between the functions $\alpha_1(S)$ and $\alpha_2(S)$:

$$\alpha_2(S) = c\ \alpha_1(S) \ , \qquad (2.3.21)$$

where c is a constant, and that for the variances along the ascending diagonal arrays we can fit the following function:

$$\mu_2(x_1|S) = bS^2 + aS \ , \qquad (2.3.22)$$

where b and a are constants.

mixed with the Generalized Inverse Gaussian distribution, or for short GIG, which is defined as:

$$\Psi(\lambda) = \frac{\left[\frac{2\sqrt{1-\theta}}{\alpha\theta}\right]^{\gamma}}{2K_{\gamma}(\alpha\sqrt{1-\theta})}\lambda^{\gamma-1} e^{-\left[\frac{1-\theta}{\theta}\right]\lambda - \frac{\alpha^2\theta}{4\lambda}}, \quad \text{where} \qquad (2.4.3)$$

$-\infty < \gamma < \infty$, $0 \leq \theta \leq 1$, $\alpha > 0$ and

$K_{\gamma}(z)$ is defined as before.

The first two moments around the origin for the GIGP are:

$$\mu_1'(S) = \frac{\alpha\theta}{2\sqrt{1-\theta}} \frac{K_{\gamma+1}(\alpha\sqrt{1-\theta})}{K_{\gamma}(\alpha\sqrt{1-\theta})} \quad \text{and} \qquad (2.4.4)$$

$$\mu_2'(S) = \frac{\alpha\theta}{2\sqrt{1-\theta}} \frac{K_{\gamma+1}(\alpha\sqrt{1-\theta})}{K_{\gamma}(\alpha\sqrt{1-\theta})} + \left[\frac{\alpha\theta}{2\sqrt{1-\theta}}\right]^2 \frac{K_{\gamma+2}(\alpha\sqrt{1-\theta})}{K_{\gamma}(\alpha\sqrt{1-\theta})} \qquad (2.4.5)$$

2.4.2 <u>Inverse Gaussian-Poisson distribution</u>

For the third marginal consider now the Inverse Gaussian-Poisson, for short IGP, which is a particular case of the GIGP if $\gamma = -1/2$. It is defined as:

$$f(S) = \sqrt{\frac{2\alpha}{\pi}}\, e^{\alpha\sqrt{1-\theta}} \frac{(\alpha\theta/2)^S}{S!} K_{S-1/2}(\alpha), \quad \text{where} \qquad (2.4.6)$$

$S = 0,1,2\ldots$

For the IGP distribution the first two moments around the origin are:

$$\mu_1'(S) = \frac{\alpha\theta}{2\sqrt{1-\theta}} \quad \text{and} \qquad (2.4.7)$$

$$\mu_2'(S) = \frac{\alpha\theta}{2\sqrt{1-\theta}} + \left[\frac{\alpha\theta}{2\sqrt{1-\theta}}\right]^2 \left[\frac{1}{\alpha\sqrt{1-\theta}} + 1\right]. \qquad (2.4.8)$$

2.4.3 Negative Binomial distribution

Consider now the Negative Binomial distribution, for short NBD, for the third marginal. It is a particular case of the GIGP when in equation (2.4.1) one takes $\gamma > 0$ and $\alpha = 0$.

In the general Bessel theory it is known that for $z << \nu$ one obtains the approximation

$$K_\nu(z) \sim \left[\frac{2}{z}\right]^\nu \frac{\Gamma(\nu)}{2} . \qquad (2.4.9)$$

In order to prove that the NBD is obtained from the GIGP one can use for $\alpha << \gamma$ the following approximations:

$$K_{S+\gamma}(\alpha) \to (2/\alpha)^{S+\gamma} \frac{\Gamma(S+\gamma)}{2} \quad \text{and} \qquad (2.4.10)$$

$$K_\gamma(\alpha\sqrt{1-\theta}) \to \left[\frac{2}{\alpha\sqrt{1-\theta}}\right]^\gamma \frac{\Gamma(\gamma)}{2} . \qquad (2.4.11)$$

This leads to

$$f(S) = \frac{(1-\theta)^\gamma}{\Gamma(\gamma)} \frac{\Gamma(S+\gamma)}{\Gamma(S+1)} \theta^S \quad \text{where} \qquad (2.4.12)$$

$S = 0,1,2,\ldots$ and $0 < \theta < 1$.

This is the standard NBD with the first two moments around the origin defined as

$$\mu_1'(S) = \frac{\gamma\theta}{1-\theta} \quad \text{and} \qquad (2.4.13)$$

$$\mu_2'(S) = \frac{\gamma\theta}{(1-\theta)^2} (1+\gamma\theta) . \qquad (2.4.14)$$

By using the ratio function for the NBD one can find that if $0 < \theta < 1$ then $0 < \gamma \leq 1$ and the distribution is reverse J-shaped.

If $\gamma > 1/\theta$ the NBD is unimodal.

2.4.3 Negative Binomial distribution

Consider now the Negative Binomial distribution, for short NBD, for the third marginal. It is a particular case of the GIGP when in equation (2.4.1) one takes $\gamma > 0$ and $\alpha = 0$.

In the general Bessel theory it is known that for $z \ll \nu$ one obtains the approximation

$$K_\nu(z) \sim \left[\frac{2}{z}\right]^\nu \frac{\Gamma(\nu)}{2} . \tag{2.4.9}$$

In order to prove that the NBD is obtained from the GIGP one can use for $\alpha \ll \gamma$ the following approximations:

$$K_{S+\gamma}(\alpha) \to (2/\alpha)^{S+\gamma} \frac{\Gamma(S+\gamma)}{2} \quad \text{and} \tag{2.4.10}$$

$$K_\gamma(\alpha\sqrt{1-\theta}) \to \left[\frac{2}{\alpha\sqrt{1-\theta}}\right]^\gamma \frac{\Gamma(\gamma)}{2} . \tag{2.4.11}$$

This leads to

$$f(S) = \frac{(1-\theta)^\gamma}{\Gamma(\gamma)} \frac{\Gamma(S+\gamma)}{\Gamma(S+1)} \theta^S \quad \text{where} \tag{2.4.12}$$

$S = 0,1,2,\ldots$ and $0 < \theta < 1$.

This is the standard NBD with the first two moments around the origin defined as

$$\mu_1'(S) = \frac{\gamma\theta}{1-\theta} \quad \text{and} \tag{2.4.13}$$

$$\mu_2'(S) = \frac{\gamma\theta}{(1-\theta)^2} (1+\gamma\theta) . \tag{2.4.14}$$

By using the ratio function for the NBD one can find that if $0 < \theta < 1$ then $0 < \gamma \leq 1$ and the distribution is reverse J-shaped.

If $\gamma > 1/\theta$ the NBD is unimodal.

2.5 Joint Distributions

This section deals with the bivariate models built from different distributions assumed for the third marginal and the ascending diagonal arrays.

2.5.1 The Beta-Binomial distribution for the ascending diagonal arrays and any distribution for the third marginal

Let us consider for the ascending diagonal arrays the BBD with the parameters S, α_1 and α_2 and for the third marginal any distribution $f(S)$ for which the first two moments around the origin are given by $\mu_1'(S)$ and $\mu_2'(S)$ and the second moment around the mean is $\mu_2(S)$.

The joint distribution of x_1 and x_2 is:

$$P_{x_1x_2}(x_1,x_2) = \begin{bmatrix} x_1+x_2 \\ x_1 \end{bmatrix} \frac{B(x_1+\alpha_1, x_2+\alpha_2)}{B(\alpha_1,\alpha_2)} f(x_1+x_2) \text{ , where} \qquad (2.5.1)$$

$S = x_1 + x_2$, $x_1 = 0,1,\ldots,S$ and $x_2 = 0,1,2,\ldots,S$.

The marginal distributions of x_1 and x_2 derived from the bivariate table are written under summation form as follows:

$$P_{x_1}(x_1) = \sum_{x_2} \begin{bmatrix} x_1+x_2 \\ x_1 \end{bmatrix} \frac{B(x_1+\alpha_1, x_2+\alpha_2)}{B(\alpha_1,\alpha_2)} f(x_1+x_2) \text{ and} \qquad (2.5.2)$$

$$P_{x_2}(x_2) = \sum_{x_1} \begin{bmatrix} x_1+x_2 \\ x_1 \end{bmatrix} \frac{B(x_1+\alpha_1, x_2+\alpha_2)}{B(\alpha_1,\alpha_2)} f(x_1+x_2) \text{ , where}$$

$S = x_1+x_2$.

Using the results of Theorem 2.1 we obtain the first moments around the origin for the marginal distributions. They are given by:

$$\mu_1'(x_1) = \frac{\alpha_1}{\alpha_1+\alpha_2}\,\mu_1'(S) \quad \text{and} \tag{2.5.3}$$

$$\mu_1'(x_2) = \mu_1'(S) - \frac{\alpha_1}{\alpha_1+\alpha_2}\,\mu_1'(S) = \frac{\alpha_2}{\alpha_1+\alpha_2}\,\mu_1'(S)\ . \tag{2.5.4}$$

In a similar way and using the results of Theorem 2.2 we obtain the second moments around the mean for the marginal distributions. They are given by:

$$\mu_2(x_1) = \frac{\alpha_1(\alpha_1+1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}\,\mu_2'(S) + \frac{\alpha_1\,\alpha_2}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}\,\mu_1'(S)$$

$$- \left[\frac{\alpha_1}{\alpha_1+\alpha_2}\right]^2 \mu_1'^2(S) \quad \text{and} \tag{2.5.5}$$

$$\mu_2(x_2) = \frac{\alpha_2(\alpha_2+1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}\,\mu_2'(S) + \frac{\alpha_1\,\alpha_2}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}\,\mu_1'(S)$$

$$- \left[\frac{\alpha_2}{\alpha_1+\alpha_2}\right]^2 \mu_1'^2(S) \tag{2.5.6}$$

or

$$\mu_2(x_2) = \mu_2(x_1) + \frac{\alpha_2-\alpha_1}{\alpha_1+\alpha_2}\,\mu_2(S)\ .$$

In order to calculate the correlation coefficient between the variables x_1 and x_2 it is necessary to calculate the first moment around the origin of the joint distribution of x_1 and x_2, and the covariance of the x_1 and x_2. By applying Theorem 2.3 to this particular model, we obtain the following formulae:

$$E(x_1x_2) = \frac{\alpha_1 \alpha_2}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)} [\mu_2'(S)-\mu_1'(S)] \quad \text{and}$$

$$\operatorname{cov}(x_1,x_2) = \frac{\alpha_1 \alpha_2}{\alpha_1+\alpha_2} \left[\frac{\mu_2'(S)-\mu_1'(S)}{\alpha_1+\alpha_2+1} - \frac{\mu_1'^2(S)}{\alpha_1+\alpha_2}\right] . \tag{2.5.7}$$

The correlation coefficient $\rho(x_1,x_2)$ between the variables x_1 and x_2 is obtained by substituting the respective expressions for $\operatorname{cov}(x_1,x_2)$, $\mu_2(x_1)$ and $\mu_2(x_2)$ into its definition

$$\rho(x_1,x_2) = \frac{\operatorname{cov}(x_1,x_2)}{\sqrt{\mu_2(x_1)\mu_2(x_2)}} .$$

The correlation coefficient is given therefore by:

$$\rho(x_1,x_2) = \sqrt{\alpha_1\alpha_2} \left[\frac{\mu_2'(S)-\mu_1'(S)}{\alpha_1+\alpha_2+1} - \frac{\mu_1'^2(S)}{\alpha_1+\alpha_2}\right]$$

$$\cdot \left[\frac{(\alpha_1+1)\mu_2'(S)+\alpha_2\mu_1'(S)}{\alpha_1+\alpha_2+1} - \frac{\alpha_1}{\alpha_1+\alpha_2}\mu_1'^2(S)\right]^{-1/2}$$

$$\cdot \left[\frac{(\alpha_2+1)\mu_2'(S)+\alpha_1\mu_1'(S)}{\alpha_1+\alpha_2+1} - \frac{\alpha_2}{\alpha_1+\alpha_2}\mu_1'^2(S)\right]^{-1/2} . \tag{2.5.8}$$

For the particular case of a symmetrical BBD, i.e. $\alpha_1 = \alpha_2$, the expressions (2.5.3) to (2.5.8) become much simpler. The following results are then obtained:

$$\mu_1'(x_1) = \mu_1'(x_2) = \frac{\mu_1'(S)}{2} , \tag{2.5.9}$$

$$\mu_2(x_1) = \mu_2(x_2) = \frac{(\alpha_1+1)\mu_2'(S)+\alpha_1\mu_1'(S)}{2(2\alpha_1+1)} - \frac{\mu_1'^2(S)}{4} , \tag{2.5.10}$$

$$E(x_1x_2) = \frac{\alpha_1}{2(2\alpha_1+1)} \left[\mu_2'(S)-\mu_1'(S)\right] , \tag{2.5.11}$$

$$\operatorname{cov}(x_1, x_2) = \frac{\alpha_1}{2}\left[\frac{\mu_2'(S)-\mu_1'(S)}{2\alpha_1+1} - \frac{\mu_1'^2(S)}{2\alpha_1}\right] \quad \text{and} \qquad (2.5.12)$$

$$\rho(x_1, x_2) = \alpha_1\left[\frac{\mu_2'(S)-\mu_1'(S)}{2\alpha_1+1} - \frac{\mu_1'^2(S)}{2\alpha_1}\right]$$

$$\cdot\left[\frac{(\alpha_1+1)\mu_2'(S)+\alpha_1\mu_1'(S)}{(2\alpha_1+1)} - \frac{\mu_1'^2(S)}{2}\right]^{-1} . \qquad (2.5.13)$$

Based on this new principle we are able to calculate the first and the second moments around the origin, the second moment around the mean for the marginal distributions and the correlation between the variables, without working through the classical approach that makes use of marginal means, variances and covariances.

In numerous situations it is difficult to develop the joint probability law for the x_1 and x_2 variables, hence it is not possible to calculate the correlations by using the classical theory. With the new approach it is not necessery to know the probability law of the marginals, and still we are able to find the correlations in a bivariate distribution.

Further we consider a particular case of the BBD which is the Binomial distribution. As mentioned above this is obtained by the limiting process of α_1 and α_2.

2.5.2 The Binomial distribution for the ascending diagonal arrays and any distribution for the third marginal

Let us assume that the variable on each ascending diagonal array is distributed according to the Binomial probability law with the same parameter p along the different arrays.

For the Binomial distribution the two moments around the origin are

$$\mu_1'(x_1|S) = Sp \text{ , and} \tag{2.5.14}$$

$$\mu_2'(x_1|S) = Sp(1-p) + S^2p^2 \text{ .} \tag{2.5.15}$$

On the third marginal we assume for the variable S any distribution $f(S)$ with the two moments around the origin $\mu_1'(S)$ and $\mu_2'(S)$ and the second central moment $\mu_2(S)$. The joint distribution of x_1 and x_2 is:

$$P(x_1,x_2) = \begin{bmatrix} x_1+x_2 \\ x_1 \end{bmatrix} p^{x_1} (1-p)^{x_2} f(x_1+x_2) \text{ .} \tag{2.5.16}$$

The marginal distributions of x_1 and x_2 derived from the bivariate table are:

$$P_{x_1}(x_1) = \frac{p^{x_1}}{x_1!} \sum_{x_2} \frac{(x_1+x_2)!}{x_2!} (1-p)^{x_2} f(x_1+x_2) \quad \text{and}$$

$$P_{x_2}(x_2) = \frac{(1-p)^{x_2}}{x_2!} \sum_{x_1} \frac{(x_1+x_2)!}{x_1!} p^{x_1} f(x_1+x_2) \text{ .} \tag{2.5.17}$$

By applying Theorems 2.1, 2.2 and 2.3 to this bivariate case, we obtain the following results:

$$\mu_1'(x_1) = p\, \mu_1'(S) \text{ ,} \tag{2.5.18}$$

$$\mu_1'(x_2) = (1-p)\, \mu_1'(S) \text{ ,} \tag{2.5.19}$$

2.5.2 The Binomial distribution for the ascending diagonal arrays and any distribution for the third marginal

Let us assume that the variable on each ascending diagonal array is distributed according to the Binomial probability law with the same parameter p along the different arrays.

For the Binomial distribution the two moments around the origin are

$$\mu_1'(x_1|S) = Sp \ , \quad \text{and} \tag{2.5.14}$$

$$\mu_2'(x_1|S) = Sp(1-p) + S^2p^2 \ . \tag{2.5.15}$$

On the third marginal we assume for the variable S any distribution f(S) with the two moments around the origin $\mu_1'(S)$ and $\mu_2'(S)$ and the second central moment $\mu_2(S)$. The joint distribution of x_1 and x_2 is:

$$P(x_1,x_2) = \begin{bmatrix} x_1+x_2 \\ x_1 \end{bmatrix} p^{x_1} (1-p)^{x_2} f(x_1+x_2) \ . \tag{2.5.16}$$

The marginal distributions of x_1 and x_2 derived from the bivariate table are:

$$P_{x_1}(x_1) = \frac{p^{x_1}}{x_1!} \sum_{x_2} \frac{(x_1+x_2)!}{x_2!} (1-p)^{x_2} f(x_1+x_2) \quad \text{and}$$

$$P_{x_2}(x_2) = \frac{(1-p)^{x_2}}{x_2!} \sum_{x_1} \frac{(x_1+x_2)!}{x_1!} p^{x_1} f(x_1+x_2) \ . \tag{2.5.17}$$

By applying Theorems 2.1, 2.2 and 2.3 to this bivariate case, we obtain the following results:

$$\mu_1'(x_1) = p\ \mu_1'(S) \ , \tag{2.5.18}$$

$$\mu_1'(x_2) = (1-p)\ \mu_1'(S) \ , \tag{2.5.19}$$

$$\mu_2(x_1) = p\ [(1-p)\ \mu_1'(S)+p\ \mu_2(S)]\ , \tag{2.5.20}$$

$$\mu_2(x_2) = (1-p)\ [p\ \mu_1'(S)+(1-p)\ \mu_2(S)]\ , \tag{2.5.21}$$

$$\mathrm{cov}(x_1,x_2) = p(1-p)[\mu_2(S)-\mu_1'(S)] \quad \text{and} \tag{2.5.22}$$

$$\rho(x_1,x_2) = \frac{\mu_2(S)-\mu_1'(S)}{\sqrt{\left[\left[\frac{1-p}{p}\right]\mu_1'(S)+\mu_2(S)\right]\left[\left[\frac{p}{1-p}\right]\mu_1'(S)+\mu_2(S)\right]}}\ . \tag{2.5.23}$$

From the above results we conclude that the parameter p of the Binomial distribution for the ascending diagonal array has no influence on the sign of the correlation coefficient. The decision whether the correlation between x_1 and x_2 is positive, negative or zero is entirely based on the parameters of the third marginal distribution.

In the special case of variables on the ascending diagonal arrays having the symmetrical Binomial distribution with $p = 1/2$, the equations (2.5.18) to (2.5.23) become:

$$\mu_1'(x_1) = \mu_1'(x_2) = \frac{\mu_1'(S)}{2}\ , \tag{2.5.24}$$

$$\mu_2(x_1) = \mu_2(x_2) = \frac{\mu_1'(S)+\mu_2(S)}{4}\ , \tag{2.5.25}$$

$$\mathrm{cov}(x_1,x_2) = \frac{\mu_2(S)-\mu_1'(S)}{4} \quad \text{and} \tag{2.5.26}$$

$$\rho(x_1,x_2) = \frac{\mu_2(S)-\mu_1'(S)}{\mu_2(S)+\mu_1'(S)}\ . \tag{2.5.27}$$

2.5.3 The Beta-Binomial distribution for the ascending diagonal arrays and the Generalized Inverse Gaussian-Poisson distribution for the third marginal

The bivariate distribution built from the GIGP for the third marginal and the BBD for the ascending diagonal arrays is

$$P(x_1,x_2) = \frac{(\sqrt{1-\theta})^{\gamma}}{K_{\gamma}(\alpha\sqrt{1-\theta})B(\alpha_1,\alpha_2)} \frac{B(x_1+\alpha_1,x_2+\alpha_2)}{x_1!x_2!} \left[\frac{\alpha\theta}{2}\right]^{x_1+x_2} K_{x_1+x_2+\gamma}(\alpha) \tag{2.5.28}$$

By applying the conventional approach to the joint distribution of x_1 and x_2 we obtain the probability law for the marginals given as

$$P_{x_1}(x_1) = \frac{(\sqrt{1-\theta})^{\gamma}}{K_{\gamma}(\alpha\sqrt{1-\theta})B(\alpha_1,\alpha_2)} \frac{\left[\frac{\alpha\theta}{2}\right]^{x_1}\Gamma(x_1+\alpha_1)}{\Gamma(x_1+1)}$$

$$\cdot \sum_{x_2} \frac{\Gamma(x_2+\alpha_2)}{\Gamma(x_2+1)\Gamma(x_1+x_2+\alpha_1+\alpha_2)} \left[\frac{\alpha\theta}{2}\right]^{x_2} K_{x_1+x_2+\gamma}(\alpha)$$

and

$$P_{x_2}(x_2) = \frac{(\sqrt{1-\theta})^{\gamma}}{K_{\gamma}(\alpha\sqrt{1-\theta})B(\alpha_1,\alpha_2)} \frac{\left[\frac{\alpha\theta}{2}\right]^{x_2}\Gamma(x_2+\alpha_2)}{\Gamma(x_2+1)}$$

$$\cdot \sum_{x_1} \frac{\Gamma(x_1+\alpha_1)}{\Gamma(x_1+1)\Gamma(x_1+x_2+\alpha_1+\alpha_2)} \left[\frac{\alpha\theta}{2}\right]^{x_1} K_{x_1+x_2+\gamma}(\alpha) \ . \tag{2.5.29}$$

These probability laws can not be expressed in a closed form and it may be that the moments, the covariance and the correlation can not be obtained from the above formulae. However, by substituting the expressions for the first two moments around the origin for the third marginal in the equations (2.5.3) to (2.5.8) we obtain the results below. For simplicity we use the substitution $u = \alpha\sqrt{1-\theta}$.