$$\begin{aligned}\operatorname{cov}(x_1,x_2) &= \sum_n \left[\sum_S S\,\mu_1'(x_1|S)\,\lambda(S|n)\right] f(n)\\ &- \sum_n \left[\sum_S \mu_2'(x_1|S)\,\lambda(S|n)\right] f(n)\\ &- \left[\sum_n \left[\sum_S \mu_1'(x_1|S)\,\lambda(S|n)\right] f(n)\right]\\ &\cdot \left[\sum_n \mu_1'(S|n)\,f(n) - \sum_n \left[\sum_S \mu_1'(x_1|S)\,\lambda(S|n)\right] f(n)\right].\end{aligned}$$

These formulae are very complex because the above theory was developed for any distributions. The expressions become much simpler and easier to calculate for a particular model fitted with specific distributions for the ascending diagonal arrays, the third marginal and the total of variables.

The correlation coefficient expression follows from the covariance formula and the variances of x_1 and x_2.

3.2 Special Distributions for the Ascending Diagonal Arrays, the Third Marginal and the Total of Variables

Let us consider the bivariate Beta-Binomial distribution with a fixed number of total purchases n in the whole product field during a specified time period. In multibrand buying the variables x_1 and x_2 represent the number of purchase occasions of the individual brands and x_0 is the "rest", i.e. all the other brands together.

The probability law of the bivariate Beta-Binomial is given by:

$$p(x_1,x_2|n) = \frac{n!}{(n-x_1-x_2)!x_1!x_2!}\ \frac{\Gamma(\alpha_0+\alpha_1+\alpha_2)}{\Gamma(\alpha_0)\Gamma(\alpha_1)\Gamma(\alpha_2)}$$

$$\cdot\ \frac{\Gamma(n-x_1-x_2+\alpha_0)\Gamma(x_1+\alpha_1)\Gamma(x_2+\alpha_2)}{\Gamma(n+\alpha_0+\alpha_1+\alpha_2)}\ , \quad \text{where} \qquad (3.2.1)$$

$x_1 = 0,1,2,\ldots,n$,

$x_2 = 0,1,2,\ldots,n$,

$S = x_1 + x_2$ and $S \leq n$.

This is a particular case of the discrete Dirichlet-Multinomial distribution. The Dirichlet-Multinomial distribution is obtained by compounding the Multinomial probability law with the Multivariate-Beta distribution.

Substituting (3.1.2) into (3.2.1) leads to the bivariate distribution for x_1 and S

$$p(x_1,S|n) = \binom{S}{x_1} \frac{B(x_1+\alpha_1,S+x_1+\alpha_2)}{B(\alpha_1,\alpha_2)} \binom{n}{S} \frac{B(S+\alpha_1+\alpha_2,n-S+\alpha_0)}{B(\alpha_1+\alpha_2,\alpha_0)}\ . \qquad (3.2.2)$$

Consequently, the joint probability law of x_1 and S can be expressed as the product

$p(x_1,S|n) = \delta(x_1|S)\ \lambda(S|n)$,

where $\delta(x_1|S)$ is the probability distribution of the ascending diagonal arrays for a given S on the bivariate table of x_1 and x_2 and $\lambda(S|n)$ is the distribution of the third marginal on the same bivariate table.

Both distributions are univariate Beta-Binomials. The variable x_1 for each ascending diagonal array is following the Beta-Binomial distribution with parameters α_1 and α_2. The variable S for the third marginal is also distributed according to the Beta-Binomial with parameters $\alpha_1+\alpha_2$ and α_0.

Let us assume that the variable n which represents the total of variables has the probability distribution defined by f(n) with the first two moments around the origin given by $\mu_1'(n)$ and $\mu_2'(n)$. In order to apply the model presented in Section 3.1 we need the following moments:

$$\mu_1'(x_1|S) = \frac{S\alpha_1}{\alpha_1+\alpha_2} , \tag{3.2.3}$$

$$\mu_2'(x_1|S) = \frac{S\alpha_1[S(\alpha_1+1)+\alpha_2]}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)} , \tag{3.2.4}$$

$$\mu_1'(S|n) = \frac{n(\alpha_1+\alpha_2)}{\alpha_1+\alpha_2+\alpha_0} \quad \text{and} \tag{3.2.5}$$

$$\mu_2'(S|n) = \frac{n(\alpha_1+\alpha_2)[n(\alpha_1+\alpha_2+1)+\alpha_0]}{(\alpha_0+\alpha_1+\alpha_2)(\alpha_0+\alpha_1+\alpha_2+1)} \tag{3.2.6}$$

By substituting the relations (3.2.3) to (3.2.6) into the equations (3.1.5) to (3.1.9) we obtain the following results:

$$\mu_1'(x_1) = \frac{\alpha_1}{\alpha_0+\alpha_1+\alpha_2}\,\mu_1'(n) , \tag{3.2.7}$$

$$\mu_1'(x_2) = \frac{\alpha_2}{\alpha_0+\alpha_1+\alpha_2}\,\mu_1'(n) , \tag{3.2.8}$$

$$\mu_2(x_1) = \frac{\alpha_1(\alpha_1+1)}{(\alpha_0+\alpha_1+\alpha_2)(\alpha_0+\alpha_1+\alpha_2+1)}\,\mu_2'(n) \tag{3.2.9}$$

$$+ \frac{\alpha_1(\alpha_0+\alpha_2)}{(\alpha_0+\alpha_1+\alpha_2)(\alpha_0+\alpha_1+\alpha_2+1)}\,\mu_1'(n) - \left[\frac{\alpha_1}{\alpha_0+\alpha_1+\alpha_2}\right]^2 \mu_1'^2(n) ,$$

$$\mu_2(x_2) = \frac{\alpha_2(\alpha_2+1)}{(\alpha_0+\alpha_1+\alpha_2)(\alpha_0+\alpha_1+\alpha_2+1)}\mu_2'(n) \tag{3.2.10}$$

$$+ \frac{\alpha_2(\alpha_0+\alpha_1)}{(\alpha_0+\alpha_1+\alpha_2)(\alpha_0+\alpha_1+\alpha_2+1)}\mu_1'(n) - \left[\frac{\alpha_2}{\alpha_0+\alpha_1+\alpha_2}\right]^2 \mu_1'^2(n) ,$$

$$\operatorname{cov}(x_1,x_2) = \frac{\alpha_1\alpha_2}{\alpha_0+\alpha_1+\alpha_2}\left[\frac{\mu_2'(n)-\mu_1'(n)}{\alpha_0+\alpha_1+\alpha_2+1} - \frac{\mu_1'^2(n)}{\alpha_0+\alpha_1+\alpha_2}\right] \tag{3.2.11}$$

and

$$\rho(x_1,x_2) = \left[\frac{\mu_2'(n)-\mu_1'(n)}{\alpha_0+\alpha_1+\alpha_2+1} - \frac{\mu_1'^2(n)}{\alpha_0+\alpha_1+\alpha_2}\right] \tag{3.2.12}$$

$$\cdot\left[\frac{\alpha_1+1}{\alpha_1(\alpha_0+\alpha_1+\alpha_2+1)}\mu_2'(n) + \frac{\alpha_0+\alpha_2}{\alpha_1(\alpha_0+\alpha_1+\alpha_2+1)}\mu_1'(n) - \frac{\mu_1'^2(n)}{\alpha_0+\alpha_1+\alpha_2}\right]^{-1/2}$$

$$\cdot\left[\frac{\alpha_2+1}{\alpha_2(\alpha_0+\alpha_1+\alpha_2+1)}\mu_2'(n) + \frac{\alpha_0+\alpha_1}{\alpha_2(\alpha_0+\alpha_1+\alpha_2+1)}\mu_1'(n) - \frac{\mu_1'^2(n)}{\alpha_0+\alpha_1+\alpha_2}\right]^{-1/2} .$$

If we assume that the total of variables is distributed in accordance with the NBD, with the first moments around the origin defined as

$$\mu_1'(n) = \frac{\gamma\theta}{1-\theta} \quad \text{and} \tag{3.2.13}$$

$$\mu_2'(n) = \frac{\gamma\theta(1+\gamma\theta)}{(1-\theta)^2} , \tag{3.2.14}$$

then we obtain the following results:

$$\mu_1'(x_1) = \frac{\alpha_1}{\alpha_0+\alpha_1+\alpha_2}\,\frac{\gamma\theta}{1-\theta} , \tag{3.2.15}$$

$$\mu_1'(x_2) = \frac{\alpha_2}{\alpha_0+\alpha_1+\alpha_2}\,\frac{\gamma\theta}{1-\theta} , \tag{3.2.16}$$

$$\mu_2(x_1) = \frac{\alpha_1}{\alpha_0+\alpha_1+\alpha_2}\,\frac{\gamma\theta}{1-\theta} \tag{3.2.17}$$

$$\cdot\left[\frac{(\alpha_1+1)(1+\gamma\theta)}{(\alpha_0+\alpha_1+\alpha_2+1)(1-\theta)} - \frac{\alpha_1\gamma\theta}{(\alpha_0+\alpha_1+\alpha_2)(1-\theta)} + \frac{\alpha_0+\alpha_2}{\alpha_0+\alpha_1+\alpha_2+1}\right],$$

$$\mu_2(x_2) = \frac{\alpha_2}{\alpha_0+\alpha_1+\alpha_2}\,\frac{\gamma\theta}{1-\theta} \tag{3.2.18}$$

$$\cdot\left[\frac{(\alpha_2+1)(1+\gamma\theta)}{(\alpha_0+\alpha_1+\alpha_2+1)(1-\theta)} - \frac{\alpha_2\gamma\theta}{(\alpha_0+\alpha_1+\alpha_2)(1-\theta)} + \frac{\alpha_0+\alpha_1}{\alpha_0+\alpha_1+\alpha_2+1}\right]$$

and

$$\mathrm{cov}(x_1,x_2) = \frac{\alpha_1\alpha_2}{(\alpha_0+\alpha_1+\alpha_2)(\alpha_0+\alpha_1+\alpha_2+1)}\,\gamma\left[\frac{\theta}{1-\theta}\right]^2\left[1-\frac{\gamma}{\alpha_0+\alpha_1+\alpha_2}\right]. \tag{3.2.19}$$

By substituting $\gamma = \alpha_0+\alpha_1+\alpha_2$ into the equation (3.2.19) we conclude that the variables x_1 and x_2 are uncorrelated. It can be seen from the probability law of x_1 and x_2 that for $\gamma = \alpha_0+\alpha_1+\alpha_2$ the variables x_1 and x_2 are distributed independently according to the NBD with the parameters α_1, θ and α_2, θ respectively.

In addition, from equation (3.2.19) it follows that for $\gamma > \alpha_0+\alpha_1+\alpha_2$ we obtain negative correlations, i.e. $\rho(x_1,x_2) < 0$; and for $\gamma < \alpha_0+\alpha_1+\alpha_2$ we obtain positive correlations, i.e. $\rho(x_1,x_2) > 0$.

From the conventional definition of the bivariate distribution it turns out that for $\gamma = \alpha_0+\alpha_1+\alpha_2$ the variables are uncorrelated, but the mathematical development becomes cumbersome for $\gamma \neq \alpha_0+\alpha_1+\alpha_2$. The model proves to be handy for obtaining the moments for the marginals and the correlations between the variables.

The above expressions for the covariance and the variances of x_1 and x_2 prove that it is possible to obtain any range of correlations in the bivariate Beta-Binomial model. The correlation coefficients depend entirely on the parameters α_0, α_1, α_2 and on the total of variables parameters γ and θ.

C H A P T E R 4

THE MULTIVARIATE CASE

In this chapter we discuss a generalization of the models presented in Chapters 2 and 3.

4.1 The General Model

The mathematical model developed in Section 3.1 for the trivariate case can be extended to the multivariate case. Let us assume k variables such that $\sum_{i=0}^{k} x_i = n$. The means, the variances of the x_i, for $i = 1,\ldots k$, and the covariances (and the respective correlations) between any of the variables x_i and x_j, for $1 \leq i < j \leq k$, can be calculated by a reduction of the multivariate model to the trivariate case. If we are interested in the variables x_i and x_j for $1 \leq i < j \leq k$, we can look at the trivariate model built by x_i, x_j and x, where $x = \sum_{l \neq i,j} x_l$, so that $x_i + x_j + x = n$.

From this point we carry on with the procedures presented in Section 3.1.

4.2 The Discrete Dirichlet-Multinomial Distribution

The Dirichlet-Multinomial distribution was used by Goodhardt, Ehrenberg and Chatfield (1984) in their multibrand purchase model. Without loss of generality we present this distribution in relation to the multibrand buying model.

Let us assume the random variables $x_0, x_1, \ldots, x_k$ such that $\sum_{i=0}^{k} x_i = n$, where n is considered to represent the whole product field. The variables $x_1, \ldots, x_k$ can be interpreted in terms of the consumption model as the respective number of units purchased by an individual household from each one of the brands $1, \ldots, k$. The variable x_0 represents the number of units purchased, which are not from the brands $1, \ldots, k$. The x_0 is considered the "rest".

We shall assume that the purchasing probability of an individual household for a brand $i = 1, \ldots, k$ is p_i and the purchasing probability for the "rest" is p_0. The above probabilities of an individual household are subject to the constraint

$$\sum_{i=0}^{k} p_i = 1 \tag{4.2.1}$$

The joint distribution of the purchases $x_0, x_1, \ldots, x_k$ of an individual household out of a total of n units for the entire product field in a specified time span is given by the Multinomial distribution:

$$P(x_0,x_1,\ldots,x_k|n) = \frac{n!}{x_0!\,x_1!\ldots x_k!}\, p_0^{x_0}\, p_1^{x_1}\ldots p_k^{x_k}, \qquad (4.2.2)$$

where $i = 0,1,\ldots,k$ and $0 \le p_i \le 1$.

Over a large number of respondents the purchase probabilities p_i for brand i will differ from household to household. Similarly, they will differ for p_0 which is the probability of not buying brands $1,2,\ldots,k$.

The joint multivariate distribution for purchase probabilities $p_0,p_1,\ldots,p_k$ may be represented by the continuous Multivariate Beta distribution, also called the Dirichlet distribution. Its probability law is

$$h(p_0,p_1,\ldots,p_k) = \frac{\Gamma(\alpha_0+\alpha_1+\ldots+\alpha_k)}{\Gamma(\alpha_0)\Gamma(\alpha_1)\ldots\Gamma(\alpha_k)}\, p_0^{\alpha_0-1}\, p_1^{\alpha_1-1}\ldots p_k^{\alpha_k-1},$$

where parameters $\alpha_i > 0$. (4.2.3)

The joint multivariate distribution for $x_0,x_1,\ldots,x_k$ purchases of the brands $1,2,\ldots,k$ and the "rest" among a large number of different households, given a fixed number of n units bought in the entire product field during a time period of specified length, is given by the discrete Dirichlet-Multinomial distribution. The compounding process of the Multinomial distribution (Equation 4.2.2) with the Dirichlet distribution (Equation 4.2.3) leads to the following probability law, called the Dirichlet-Multinomial distribution:

$$P(x_1,\ldots,x_k|n) = \frac{n!}{(n-x_1-\ldots x_k)!x_1!\ldots x_k!}\,\frac{\Gamma(\alpha_0+\alpha_1+\ldots+\alpha_k)}{\Gamma(\alpha_0)\Gamma(\alpha_1)\ldots\Gamma(\alpha_k)}$$

$$\cdot\,\frac{\Gamma(n-x_1-\ldots-x_k+\alpha_0)\,\Gamma(x_1+\alpha_1)\ldots\Gamma(x_k+\alpha_k)}{\Gamma(n+\alpha_0+\alpha_1+\ldots+\alpha_k)}, \quad (4.2.4)$$

where $x_i = 0,1,\ldots,n.\ \forall\ i = 1,\ldots,k.$

If we are interested in the correlations between any two variables x_i and x_j, we reduce our multivariate model to a trivariate model. In fact, it is necessary to group all the other variables into a new variable $x = \sum_{\substack{l=0\\ l\neq i,j}}^{k} x_l$. The joint distribution of any two variables x_i and x_j is therefore given by:

$$P(x_i,x_j|n) = \frac{n!}{(n-x_i-x_j)!x_i!x_j!}\,\frac{\Gamma(\alpha_0+\alpha_1+\ldots+\alpha_k)}{\Gamma(\alpha_i)\,\Gamma(\alpha_j)\,\Gamma\left[\sum_{\substack{l=0\\ l\neq i,j}}^{k}\alpha_l\right]}$$

$$\cdot\,\frac{\Gamma(x_i+\alpha_i)\,\Gamma(x_j+\alpha_j)\,\Gamma\left[n-x_i-x_j+\sum_{\substack{l=0\\ l\neq i,j}}^{k}\alpha_l\right]}{\Gamma(n+\alpha_0+\alpha_1+\ldots+\alpha_k)}. \quad (4.2.5)$$

Mosimann (1962) proved that the correlations in the Dirichlet-Multinomial distribution and the correlations between any two variables in the multivariate Beta distribution (Dirichlet distribution) are the same.

For the k variate case we have the following moments of the multivariate Beta distribution:

$$E(P_i) = \frac{\alpha_i}{\sum_{i=0}^{k}\alpha_i}, \quad (4.2.6)$$

$$E(P_i^2) = \frac{\alpha_i(\alpha_i+1)}{\left[\sum_{i=0}^{k} \alpha_i+1\right]\left[\sum_{i=0}^{k} \alpha_i\right]} \text{ and} \tag{4.2.7}$$

$$E(P_i P_j) = \frac{\alpha_i \alpha_j}{\left[\sum_{i=0}^{k} \alpha_i+1\right]\left[\sum_{i=0}^{k} \alpha_i\right]} . \tag{4.2.8}$$

The corresponding moments for the Dirichlet-Multinomial distribution are:

$$E(x_i) = \frac{n\,\alpha_i}{\sum_{i=0}^{k} \alpha_i} , \tag{4.2.9}$$

$$E(x_i^2) = \frac{n(n-1)\,\alpha_i(\alpha_i+1)}{\left[\sum_{i=0}^{k} \alpha_i+1\right]\left[\sum_{i=0}^{k} \alpha_i\right]} + \frac{n\,\alpha_i}{\sum_{i=0}^{k} \alpha_i} \text{ and} \tag{4.2.10}$$

$$E(x_i x_j) = \frac{n(n-1)\,\alpha_i \alpha_j}{\left[\sum_{i=0}^{k} \alpha_i+1\right]\left[\sum_{i=0}^{k} \alpha_i\right]} . \tag{4.2.11}$$

From the above moments we calculate the variances and covariances in terms of the moments of the multivariate Beta distribution as follows:

$$\mathrm{var}(x_i) = n\,E(p_i)\left[1-E(p_i)\right]\left[\frac{n+\sum_{i=0}^{k} \alpha_i}{1+\sum_{i=0}^{k} \alpha_i}\right] , \tag{4.2.12}$$

$$\mathrm{cov}(x_i x_j) = -n\,E(p_i)E(p_j)\left[\frac{n+\sum_{i=0}^{k} \alpha_i}{1+\sum_{i=0}^{k} \alpha_i}\right] . \tag{4.2.13}$$

Therefore, for a fixed n the correlations between any two variables in the Dirichlet-Multinomial distribution are:

$$\rho_{ij} = -\left[\frac{E(p_i)E(p_j)}{[1-E(p_i)][1-E(p_j)]}\right]^{1/2} . \tag{4.2.14}$$

The same formula for correlations between any two variables is obtained in the Multivariate Beta distribution.

In Equation (4.2.5), by substituting $S = x_i + x_j$ we obtain the bivariate distribution of x_i and S:

$$P(x_i, S|n) = \left[\binom{S}{x_i}\frac{B(x_i+\alpha_i,\ S-x_i+\alpha_i)}{B(\alpha_i,\ \alpha_j)}\right]\left[\binom{n}{S}\frac{B\left[S+\alpha_i+\alpha_j,\ n-S+\sum\limits_{\substack{l=0\\ l\neq i,j}}^{k}\alpha_l\right]}{B\left[\alpha_i+\alpha_j,\ \sum\limits_{\substack{l=0\\ l\neq i,j}}^{k}\alpha_l\right]}\right]. \tag{4.2.15}$$

This is the joint distribution between x_i, the variable on the ascending diagonal array of the bivariate table of x_i and x_j and the variable S on the third marginal for the same bivariate table. In order to calculate the means, the variances and the correlations between any x_i, x_j we follow the same steps as in the trivariate model.

Let us assume that the distribution for the whole product field is given by the probability law F(n) with the first two moments around the origin $\mu_1'(n)$ and $\mu_2'(n)$.

We can denote our model as follows:

$$M(n|\underline{p}) \underset{\underline{p}}{\wedge} D(\underline{p}|\underline{\alpha}) \underset{n}{\wedge} F(n|\underline{\xi}) . \tag{4.2.16}$$

Applying the model from Section 3.1 we obtain the following results:

$$\mu_1'(x_i) = \frac{\alpha_i}{\sum_{m=0}^{k} \alpha_m} \mu_1'(n) , \tag{4.2.17}$$

$$\mu_1'(x_j) = \frac{\alpha_j}{\sum_{m=0}^{k} \alpha_m} \mu_1'(n) , \tag{4.2.18}$$

$$\mu_2(x_i) = \frac{\alpha_i(\alpha_i+1)}{\left[\sum_{m=0}^{k} \alpha_m\right]\left[\sum_{m=0}^{k} \alpha_m+1\right]} \mu_2'(n) \tag{4.2.19}$$

$$+ \frac{\alpha_i \left[\sum_{\substack{l=0 \\ l\neq i}}^{k} \alpha_l\right]}{\left[\sum_{m=0}^{k} \alpha_m\right]\left[\sum_{m=0}^{k} \alpha_m+1\right]} \mu_1'(n) - \left[\frac{\alpha_i}{\sum_{m=0}^{k} \alpha_m}\right]^{1/2} \mu_1'^2(n) ,$$

$$\mu_2(x_j) = \frac{\alpha_j(\alpha_j+1)}{\left[\sum_{m=0}^{k} \alpha_m\right]\left[\sum_{m=0}^{k} \alpha_m+1\right]} \mu_2'(n) \tag{4.2.20}$$

$$+ \frac{\alpha_j \left[\sum_{\substack{l=0 \\ l\neq j}}^{k} \alpha_l\right]}{\left[\sum_{m=0}^{k} \alpha_m\right]\left[\sum_{m=0}^{k} \alpha_m+1\right]} \mu_1'(n) - \left[\frac{\alpha_j}{\sum_{m=0}^{k} \alpha_m}\right]^{1/2} \mu_1'^2(n) ,$$

$$E(x_i x_j) = \frac{\alpha_i \ \alpha_j}{\left[\sum_{m=0}^{k} \alpha_m\right]\left[\sum_{m=0}^{k} \alpha_m+1\right]} \left[\mu_2'(n) - \mu_1'(n)\right] \quad \text{and} \tag{4.2.21}$$

$$\mathrm{cov}(x_i, x_j) = \frac{\alpha_i \alpha_j}{\sum_{m=0}^{k} \alpha_m} \left[\frac{\mu_2'(n)-\mu_1'(n)}{\sum_{m=0}^{k} \alpha_m+1} - \frac{\mu_1'^2(n)}{\sum_{m=0}^{k} \alpha_m}\right] . \tag{4.2.22}$$

Expressed in terms of the Multivariate Beta distribution, the moments of Equations (4.2.17) to (4.2.22) lead to the following expressions:

$$\mu_1'(x_i) = E(p_i)\ \mu_1'(n) \tag{4.2.23}$$

$$\mu_1'(x_j) = E(p_j)\ \mu_1'(n) \tag{4.2.24}$$

$$\mu_2(x_i) = E(p_i^2)[\mu_2'(n)-\mu_1'(n)] + E(p_i)[1-E(p_i)\mu_1'(n)]\mu_1'(n) \tag{4.2.25}$$

$$\mu_2(x_j) = E(p_j^2)[\mu_2'(n)-\mu_1'(n)] + E(p_j)[1-E(p_j)\mu_1'(n)]\mu_1'(n) \tag{4.2.26}$$

$$\mathrm{cov}(x_i,x_j) = E(p_ip_j)[\mu_2'(n)-\mu_1'(n)] - E(p_i)E(p_j)\mu_1'^2(n) \tag{4.2.27}$$

The correlation ρ_{ij} between the variables x_i and x_j is therefore given by:

$$\rho_{ij} = \left[E(p_ip_j) - E(p_i)E(p_j)\frac{\mu_1'^2(n)}{\mu_2'(n)-\mu_1'(n)}\right] \cdot \left[E(p_i^2) + E(p_i)\left[1-E(p_i)\mu_1'(n)\right]\frac{\mu_1'(n)}{\mu_2'(n)-\mu_1'(n)}\right]^{-1/2} \cdot \left[E(p_j^2) + E(p_j)\left[1-E(p_j)\mu_1'(n)\right]\frac{\mu_1'(n)}{\mu_2'(n)-\mu_1'(n)}\right]^{-1/2} . \tag{4.2.28}$$

The correlations are different from the ones obtained from the Dirichlet-Multinomial distribution for a fixed n. By allowing n to follow any distribution we obtain a range of correlations that is wider than the correlations obtained in the Multivariate Beta distribution.

If we assume that the distribution for the whole product field n is the NBD, which is a function of the parameters γ and θ, then we obtain the buying behaviour model presented by Goodhardt, Ehrenberg and Chatfield (1984).

The advantage of the NBD-Dirichlet model presented in the above form in terms of the ascending diagonal arrays and the third marginal for each bivariate distribution of x_i, x_j and the whole product field is that it provides an easy way to calculate the means and the variances of the marginal variables.

The above model enables the calculations of the means, variance-covariance matrix and the correlations between any two brands without calculating the marginal distributions. As mentioned by Goodhardt, Ehrenberg and Chatfield (1984), it is possible to express the marginal distributions in a closed form by using the four-parameter hypergeometric function as presented by Abramowitz and Stegun (1965). Since the marginal distributions must be evaluated through infinite summations, it is obvious that one will find great difficulty in calculating the means, variances and correlations.

As shown in Section 2.5.5 one can obtain a relatively wide range of correlations between any two brands. In the multivariate case one can also obtain a relatively large range of correlations between any two brands, by reducing the model to the trivariate case. From Equation (4.2.22) one obtains negative correlations if:

$$\frac{\mu_2'(n)-\mu_1'(n)}{\sum\limits_{m=0}^{k} \alpha_m+1} - \frac{\mu_1'^2(n)}{\sum\limits_{m=0}^{k} \alpha_m} < 0 \qquad (4.2.29)$$

and positive correlations if:

$$\frac{\mu_2'(n)-\mu_1'(n)}{\sum\limits_{m=0}^{k} \alpha_m+1} - \frac{\mu_1'^2(n)}{\sum\limits_{m=0}^{k} \alpha_m} > 0 \tag{4.2.30}$$

If we assume the NBD-Dirichlet model then the first and second moments around the origin are

$$\mu_1'(n) = \frac{\gamma\theta}{1-\theta} \quad \text{and} \quad \mu_2' = \frac{\gamma\theta(1+\gamma\theta)}{(1-\theta)^2} \tag{4.2.31}$$

and by substituting the above relations into Equation (4.2.22) one obtains positive correlations for

$$\sum_{m=0}^{k} \alpha_m > \gamma \tag{4.2.32}$$

and negative correlations for

$$\sum_{m=0}^{k} \alpha_m < \gamma \; . \tag{4.2.33}$$

The following examples are presented in order to emphasize the above theory.

Example 1

Assume a symmetrical case of four brands and the "rest" with the following parameters: $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$ and assume for the whole product field an NBD with the parameters $\gamma = 0,1$ and $\theta = 0,95$. The correlation coefficient obtained between any two brands is $\rho_{ij} = 0,409$.

One can obtain a considerable correlation even for a higher number of brands as shown in Example 2.

Example 2

Assume the same distributions as in Example 1 with the parameters $\alpha_0 = \alpha_1 = \ldots = \alpha_9 = 1$, $\gamma = 0,1$, and $\theta = 0,95$. The correlation coefficient obtained is $\rho_{ij} = 0,371$.

Example 3

One can obtain also negative correlations as seen in Equation (4.2.18). Assume the same distributions as in Example 1 with the parameters $\alpha_0 = \ldots = \alpha_4 = 1$, $\gamma = 16$ and $\theta = 0,8$. The correlation coefficient obtained is $\rho_{ij} = -0,135$.

The correlations between any two brands vary in absolute value as a function of $\left|\sum_{m=0}^{k} \alpha_m - \gamma\right|$ and θ.

By assuming that n is fixed in Examples 1 and 2 we deal with the Dirichlet-Multinomial distribution and the correlations between any two variables are respectively -0,25 and -0,125. These are negative correlations and are much lower in absolute value compared to the ones obtained by taking the NBD as $F(n|\gamma,\theta)$ in equation (4.2.16).

The parameters $\alpha_0, \alpha_1, \ldots, \alpha_k$ can be estimated through the third marginal and the ascending diagonal arrays for each bivariate table of x_i, x_j. The parameters γ and θ for the NBD will be estimated from the data for the whole product field. The methods of estimation for the above parameters will depend entirely on the type of data.

The model presented in Equation (4.2.4) assumes a fixed number of units bought from the entire product field. One can assume any suitable distribution for the whole product field, not only the NBD. In this way the above mentioned model opens new perspectives for further work in the multibrand purchase field and in general in compositional data.

CHAPTER 5

GENERAL PROOFS OF POISSON MIXTURES

In the previous chapters we used several Compound Poisson distributions for the third marginal. In this chapter we investigate the general properties of the univariate and the bivariate case of the compound Poisson distributions. We are concerned with the relations between the moments and the various statistical indices such as coefficients of dispersion, variation, skewness and kurtosis of the compound Poisson and the mixing distribution.

In general the compound Poisson moments will be expressed as functions of the exposure period t and the moments of the observational period for $t = 1$.

5.1 The Univariate Case

Most of the distributions which are applied in the general models dealt with in this investigation are Poisson mixtures. An analysis of the Poisson mixtures and their properties follows.

A compound Poisson distribution is defined as:

$$f(x|t) = \int_0^\infty p(x|t,\lambda)\ \Psi(\lambda)\ d\lambda\ , \tag{5.1.1}$$

where $$p(x|t,\lambda) = \frac{e^{-t\lambda}\ (t\lambda)^x}{x!} \quad \text{for}\quad x = 0,1,2,\ldots\ . \tag{5.1.2}$$

The parameter λ refers to the unit length of exposure and t is a multiple of this unit. The distribution of λ is given by $\Psi(\lambda)$, which is a continuous probability function, also called the mixing distribution.

Property 5.1

The following relations exist between the first moment around the origin and the second, third and fourth central moments of the compound Poisson and the mixing distribution:

$$\mu_1'(x|t) = t\ \mu_1'(\lambda)\ , \tag{5.1.3}$$

$$\mu_2(x|t) = t^2\mu_2(\lambda) + t\ \mu_1'(\lambda)\ , \tag{5.1.4}$$

$$\mu_3(x|t) = t^3\mu_3(\lambda) + 3t^2\mu_2(\lambda) + t\ \mu_1'(\lambda) \quad \text{and} \tag{5.1.5}$$

$$\mu_4(x|t) = t^4\mu_4(\lambda) + 6t^3\mu_3(\lambda) + 7t^2\mu_2(\lambda) + (3t^2+t)\mu_1'(\lambda). \tag{5.1.6}$$

Proof

From the definition of the first moment around the origin one obtains:

$$\begin{aligned}\mu_1'(x|t) &= \sum_{x=0}^{\infty} \int_0^\infty x\ \frac{e^{-t\lambda}\ (t\lambda)^x}{x!}\ \Psi(\lambda)\ d\lambda \\ &= t\int_0^\infty \lambda\ \Psi(\lambda)\ d\lambda = t\ \mu_1'(\lambda)\ .\end{aligned}$$

Similarly, for the second moment around the origin we have:

$$\mu_2'(x|t) = \sum_{x=0}^{\infty} \int_0^{\infty} x^2 \frac{e^{-t\lambda}(t\lambda)^x}{x!} \Psi(\lambda)\, d\lambda$$

$$= \int_0^{\infty} (t\lambda + t^2\lambda^2)\, \Psi(\lambda)\, d\lambda = t\, \mu_1'(\lambda) + t^2\, \mu_2'(\lambda)\ ,$$

and from the definition of $\mu_2(x|t)$

$$\mu_2(x|t) = \mu_2'(x|t) - \mu_1'^2(x|t)$$

we conclude that

$$\mu_2(x|t) = t^2\mu_2(\lambda) + t\mu_1'(\lambda)\ .$$

The same procedure applies for the third and fourth central moments.

$$\mu_3'(x|t) = \sum_{x=0}^{\infty} \int_0^{\infty} x^3 \frac{e^{-t\lambda}(t\lambda)^x}{x!} \Psi(\lambda)\, d\lambda$$

$$= t^3\mu_3'(\lambda) + 3t^2\mu_2'(\lambda) + t\mu_1'(\lambda)$$

concluding that

$$\mu_3(x|t) = t^3\mu_3(\lambda) + 3t^2\mu_2(\lambda) + t\mu_1'(\lambda)\ ,$$

and

$$\mu_4'(x|t) = \sum_{x=0}^{\infty} \int_0^{\infty} x^4 \frac{e^{-t\lambda}(t\lambda)^x}{x!} \Psi(\lambda)\, d\lambda$$

$$= t^4\mu_4'(\lambda) + 6t^3\mu_3'(\lambda) + 7t^2\mu_2'(\lambda) + t\mu_1'(\lambda)\ .$$

concluding that

$$\mu_4(x|t) = t^4\mu_4(\lambda) + 6t^3\mu_3(\lambda) + 7t^2\mu_2(\lambda) + (3t^2+t)\mu_1'(\lambda)\ .$$

<u>Property 5.2</u>

The following relations exist between the first moment around the origin and the second, third and fourth central moments as a function of a period of length t and the respective moments at the observational period $t = 1$:

$$\mu_1'(x|t) = t\, \mu_1'(x|1) \ ,$$

$$\mu_2(x|t) = t^2\mu_2(x|1) - t(t-1)\mu_1'(x|1) \ ,$$

$$\mu_3(x|t) = t^3\mu_3(x|1) - 3t^2(t-1)\mu_2(x|1) + t(t-1)(2t-1)\mu_1'(x|1) \quad \text{and}$$

$$\mu_4(x|t) = t^4\mu_4(x|1) - 6t^3(t-1)\mu_3(x|1) + t^2(t-1)(11t-7)\mu_2(x|1) - t(t-1)(9t^2-3t+1)\mu_1'(x|1) \ .$$

Proof

These relations are easily proved based on equations (5.1.3) to (5.1.6) above.

Conclusion 1

In equation (5.1.4) for the variance of the compound Poisson distribution the variance of the mixing plays the major role when t is very large or $t \to \infty$.

In the equation $\mu_2(x|t) = t^2\mu_2(\lambda) + t\mu_1'(\lambda)$ the variance of the compound Poisson consists of two components:

$t\, \mu_1'(\lambda)$, which is the average of the Poisson variances, and

$t^2\mu_2(\lambda)$, which is the variance of the mixing distribution.

The ratio of variances of the Poisson and the mixing distribution is

$$\frac{\mu_2(x|t)}{t^2\ \mu_2(\lambda)} \ . \qquad (5.1.7)$$

In the limiting process of $t \to \infty$ we obtain

$$\lim_{t\to\infty} \frac{\mu_2(x|t)}{t^2\,\mu_2(\lambda)} = \lim_{t\to\infty} \frac{\mu_1(\lambda)}{t\,\mu_2(\lambda)} + 1 = 1 .$$

We conclude that if $t \to \infty$, then:

(a) the variance component, which is the contribution of the mixing distribution, is the dominant one; and

(b) the variance component, which is the average of the Poisson variances, can be neglected.

Conclusion 2

The expression for the variance of the compound Poisson resembles the concept contained in the analysis of variance. The following equation exists in a one-way design between the sum of squares.

SST = SSB + SSW, where

SST is the Total Sum of Squares,

SSB is the Sum of Squares Betw[illegible]e treatments and

SSW is the Sum of Squares Within treatments.

In our case:

$t^2\mu_2(\lambda)$ can be looked upon as the variance between treatments (Poisson distributions); and

$t\mu_1'(\lambda)$ can be regarded as the variance within treatments (Poisson distributions).

Conclusion 3

The third central moment for the compound Poisson $\mu_3(x|t)$ is explained in terms of the contribution of the Poisson and the mixing distributions as follows:

$\mu_3(x|t) = t^3\mu_3$ (Mixing)

+ t (Average of the third central moment of the Poisson)

+ $3t^2\mu_2$ (Due to Poisson and Mixing).

5.1.1 The coefficient of dispersion

The coefficient of dispersion for the compound Poisson is defined as:

$$\mathrm{COD}(x|t) = \frac{\mu_2(x|t)}{\mu_1'(x|t)} . \qquad (5.1.8)$$

The following expression for $\mathrm{COD}(x|t)$ is obtained by substituting the relations (5.1.3) and (5.1.4) into the equation (5.1.8):

$$\mathrm{COD}(x|t) = 1 + t\ \mathrm{COD}(\lambda), \quad \text{where} \qquad (5.1.9)$$

$\mathrm{COD}(\lambda) = \dfrac{\mu_2(\lambda)}{\mu_1'(\lambda)}$ is the coefficient of dispersion of the mixing distribution.

Due to the fact that the coefficient of dispersion for the Poisson distribution is unity, equation (5.1.9) can be written as:

$$\mathrm{COD}(x|t) = \mathrm{COD}(\text{Poisson}) + t\ \mathrm{COD}(\text{Mixing}). \qquad (5.1.10)$$

In the limiting process the coefficient c dispersion tends to ∞ for $t \to \infty$.

5.1.2 The coefficient of variation

The coefficient of variation for the compound Poisson is defined as

$$CV(x|t) = \frac{[\mu_2(x|t)]^{1/2}}{\mu_1'(x|t)} . \tag{5.1.11}$$

The following expression for $CV(x|t)$ is obtained by substituting relations (5.1.3) and (5.1.4) into equation (5.1.11):

$$CV(x|t) = \left[\frac{1}{t\mu_1'(\lambda)} + \frac{\mu_2(\lambda)}{\mu_1'^2(\lambda)}\right]^{1/2} . \tag{5.1.12}$$

The coefficient of variation for the compound Poisson approaches the coefficient of variation for the mixing distribution in the limiting process for $t \to \infty$:

$$\lim_{t\to\infty} CV(x|t) = CV(\lambda), \quad \text{where} \quad CV(\lambda) = \frac{[\mu_2(\lambda)]^{1/2}}{\mu_1'(\lambda)} .$$

Note

It is possible to express the coefficient of variation for the compound Poisson in terms of the coefficient of variation for the Poisson distribution and the coefficient of variation for the mixing distribution as follows:

$$CV(x|t) = \left[\frac{1}{t\ \mu_1'[CV^{-2}(\text{Poisson})]} + CV^2(\text{Mixing})\right]^{1/2} , \qquad (5.1.13)$$

where $CV(\text{Poisson}) = \lambda^{-1/2}$ and

$$CV(\text{Mixing}) = \frac{[\mu_2(\lambda)]^{1/2}}{\mu_1'(\lambda)} .$$

5.1.3 The coefficients of skewness and kurtosis

Theorem 5.1

The coefficient of skewness for the compound Poisson $\sqrt{\beta_1(x|t)}$ approaches the coefficient of skewness of the mixing distribution $\sqrt{\beta_1(\lambda)}$ when $t \to \infty$.

Proof

The coefficient of skewness for the compound Poisson is defined as:

$$\sqrt{\beta_1(x|t)} = \frac{\mu_3(x|t)}{[\mu_2(x|t)]^{3/2}} . \qquad (5.1.14)$$

By substituting the relations (5.1.4) and (5.1.5) into the equation (5.1.14) the coefficient of skewness can be written as:

$$\sqrt{\beta_1(x|t)} = \frac{t^3\mu_3(\lambda) + 3t^2\mu_2(\lambda) + t\mu_1'(\lambda)}{[t^2\mu_2(\lambda) + \mu_1'(\lambda)]^{3/2}} . \qquad (5.1.15)$$

The ratio between the coefficients of skewness for the compound Poisson and the mixing distribution is given by:

$$R_{Sk}(t) = \frac{\sqrt{\beta_1(x|t)}}{\sqrt{\beta_1(\lambda)}} = \frac{1 + \dfrac{3}{t\sqrt{\mu_2(\lambda)\beta_1(\lambda)}} + \dfrac{\mu_1'(\lambda)}{t^2\mu_2^{3/2}(\lambda)\sqrt{\beta_1(\lambda)}}}{\left[1 + \dfrac{\mu_1'(\lambda)}{t\mu_2(\lambda)}\right]^{3/2}} \, ,$$

where $\sqrt{\beta_1(\lambda)} = \dfrac{\mu_3(\lambda)}{[\mu_2(\lambda)]^{3/2}}$. (5.1.16)

By limiting this ratio for $t \to \infty$ we obtain

$$\lim_{t\to\infty} R_{Sk}(t) = 1.$$

We conclude, therefore, that the coefficient of skewness for the compound Poisson approaches the coefficient of skewness for the mixing distribution if $t \to \infty$.

<u>**Theorem 5.2**</u>

The coefficient of kurtosis for the compound Poisson $\beta_2(x|t)$ approaches the coefficient of kurtosis for the mixing distribution $\beta_2(\lambda)$ if $t \to \infty$.

Proof

The coefficient of kurtosis for the compound Poisson is defined as:

$$\beta_2(x|t) = \frac{\mu_4(x|t)}{\mu_2^{\,2}(x|t)} \, . \tag{5.1.17}$$

By substituting (5.1.4) and (5.1.6) into (5.1.17), the coefficient of kurtosis can be written as:

$$\beta_2(x|t) = \frac{t^4\mu_4(\lambda) + 6t^3\mu_3(\lambda) + 7t^2\mu_2(\lambda) + (3t^2+t)\mu_1'(\lambda)}{[t^2\mu_2(\lambda) + t\mu_1'(\lambda)]^2} . \quad (5.1.18)$$

The ratio between the coefficients of kurtosis for the compound Poisson and the mixing distribution is given by:

$$R_K(t) = \frac{\beta_2(x|t)}{\beta_2(\lambda)} = \frac{1 + \frac{6\mu_3(\lambda)}{t\mu_4(\lambda)} + \frac{7\mu_2(\lambda)}{t^2\mu_4(\lambda)} + \left[\frac{3}{t^2} + \frac{1}{t^3}\right]\frac{\mu_1'(\lambda)}{\mu_4(\lambda)}}{\left[1 + \frac{\mu_1'(\lambda)}{t\mu_2^2(\lambda)}\right]^2} , \quad (5.1.19)$$

where $\beta_2(\lambda) = \frac{\mu_4(\lambda)}{\mu_2^2(\lambda)}$.

By limiting this ratio for $t \to \infty$ we obtain

$\lim_{t\to\infty} R_K(t) = 1.$

We conclude, therefore, that the coefficient of kurtosis for the compound Poisson approaches the coefficient of kurtosis for the mixing distribution when $t \to \infty$.

As $t \to \infty$ the compound Poisson tends towards the mixing distribution. This result is quite surprising, because one could expect the tendency to normality.

5.1.4 Yule characteristic

Property

For all compound Poisson distributions the "Yule characteristic" (1944) can be defined in terms of the first and the second moments around the origin of the mixing distribution as follows:

$$K^* = \frac{10^4}{t}\left[\frac{\mu_2'(x|t)}{\mu_1'(x|t)} - 1\right] = 10^4 \frac{\mu_2'(\lambda)}{\mu_1'(\lambda)} . \tag{5.1.20}$$

Proof

The Yule characteristic for the compound Poisson is defined as

$$K^* = \frac{10^4}{t}\left[\frac{\mu_2'(x|t)}{\mu_1'(x|t)} - 1\right] . \tag{5.1.21}$$

By substituting (5.1.3) and (5.1.4) into (5.1.21) we obtain the following relation:

$$K^* = \frac{10^4}{t}\left[\frac{t^2\mu_2'(\lambda) + t\mu_1'(\lambda)}{t\mu_1'(x|t)} - 1\right] = 10^4 \frac{\mu_2'(\lambda)}{\mu_1'(\lambda)} .$$

5.2 The Bivariate Case

In this section we investigate some properties of the bivariate Compound Poisson.

The uncorrelated bivariate Poisson distribution with marginal means $E(x_1|t_1) = t_1\lambda$ and $E(x_2|t_2) = t_2\lambda$,

where the parameter λ refers to the unit length of exposure, and t_1 and t_2 are multiples of this unit, is:

$$P(x_1,x_2|\lambda,t_1,t_2) = \frac{e^{-(t_1+t_2)\lambda}(t_1\lambda)^{x_1}(t_2\lambda)^{x_2}}{x_1!x_2!}, \quad \text{for} \qquad (5.2.1)$$

$x_1 = 0,1,2,\ldots,\infty$,

$x_2 = 0,1,2,\ldots,\infty$ and

$t_1 > 0, \quad t_2 > 0, \quad \lambda > 0.$

If the mixing distribution is $\Psi(\lambda)$, then the resulting bivariate Poisson mixture is:

$$f(x_1,x_2|t_1t_2) = \int_0^\infty P(x_1,x_2|\lambda,t_1,t_2)\,\Psi(\lambda)\,d\lambda. \qquad (5.2.2)$$

Lemma 5.1 deals with the shape of the marginal distributions derived from the bivariate compound Poisson.

__Lemma 5.1__

The marginal distributions of the bivariate Poisson mixture are univariate compound Poisson.

Proof

The marginal distribution of x_1 is:

$$f(x_1|t_1) = \sum_{x_2=0}^{\infty} \int_0^{\infty} e^{-(t_1+t_2)\lambda} \frac{(t_1\lambda)^{x_1}(t_2\lambda)^{x_2}}{x_1!x_2!} \Psi(\lambda)\, d\lambda$$

$$= \int_0^{\infty} e^{-t_1\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \left[\sum_{x_2=0}^{\infty} e^{-t_2\lambda} \frac{(t_2\lambda)^{x_2}}{x_2!} \right] \Psi(\lambda)\, d\lambda$$

$$= \int_0^{\infty} e^{-t_1\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \Psi(\lambda)\, d\lambda \qquad (5.2.3)$$

The marginal distribution of x_1 is therefore a univariate Poisson mixture as given in definition (5.1.1).

In a similar way it can be proved that the marginal distribution of x_2

$$f(x_2|t_2) = \int_0^{\infty} e^{-t_2\lambda} \frac{(t_2\lambda)^{x_2}}{x_2!} \Psi(\lambda)\, d\lambda \qquad (5.2.4)$$

is a Poisson mixture.

Lemma 5.2 deals with the expected value of the joint distribution of x_1 and x_2.

Lemma 5.2

The expected value of x_1 and x_2 can be expressed in terms of the second moment around the origin of the mixing distribution as follows:

$$E(x_1x_2|t_1,t_2) = t_1t_2\mu_2'(\lambda) \qquad (5.2.5)$$

Proof

$$E(x_1x_2|t_1,t_2) = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} x_1x_2\int_0^{\infty} \frac{e^{-(t_1+t_2)\lambda}(t_1\lambda)^{x_1}(t_2\lambda)^{x_2}}{x_1!x_2!}\,\varphi(\lambda)\,d\lambda$$

$$= \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty}\int_0^{\infty} \frac{e^{-(t_1+t_2)\lambda}(t_1\lambda)^{x_1}(t_2\lambda)^{x_2}}{(x_1-1)!(x_2-1)!}\,\varphi(\lambda)\,d\lambda.$$

By substituting $i = x_1-1$ and $j = x_2-1$ one obtains:

$$E(x_1x_2|t_1,t_2) = \int_0^{\infty}(t_1\lambda)(t_2\lambda)\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{e^{-(t_1+t_2)\lambda}(t_1\lambda)^{i}(t_2\lambda)^{j}}{i!j!}\right]\varphi(\lambda)d\lambda$$

$$= t_1t_2\int_0^{\infty}\lambda^2\,\varphi(\lambda)\,d\lambda$$

$$= t_1t_2\,\mu_2'(\lambda).$$

Lemma 5.3

The covariance of x_1 and x_2 can be expressed in terms of the variance of the mixing distribution as follows:

$$\mathrm{cov}(x_1x_2|t_1t_2) = t_1t_2\mu_2(\lambda). \qquad (5.2.6)$$

Proof

The definition of the covariance of x_1 and x_2 is

$$\mathrm{cov}(x_1,x_2|t_1,t_2) = E(x_1x_2|t_1,t_2) - E(x_1|t_1)E(x_2|t_2)\ . \qquad (5.2.7)$$

Substitution of (5.1.3) and (5.2.5) leads to:

$$\mathrm{cov}(x_1,x_2|t_1,t_2) = t_1 t_2\left[\mu_2'(\lambda) - \mu_1'^2(\lambda)\right]$$
$$= t_1 t_2\ \mu_2(\lambda)\ .$$

Lemma 5.4

The correlation coefficient between x_1 and x_2 can be expressed in terms of the coefficient of dispersion of the mixing distribution as follows:

$$\rho(x_1,x_2|t_1,t_2) = \left[1 + \frac{1}{t_1\mathrm{COD}(\lambda)}\right]^{-1/2}\left[1 + \frac{1}{t_2\mathrm{COD}(\lambda)}\right]^{-1/2} \qquad (5.2.8)$$

where $\mathrm{COD}(\lambda) = \dfrac{\mu_2(\lambda)}{\mu_1'(\lambda)}$.

Proof

The definition of the correlation coefficient is:

$$\rho(x_1,x_2|t_1,t_2) = \frac{\mathrm{cov}(x_1,x_2|t_1,t_2)}{\sqrt{\mu_2(x_1|t_1)\ \mu_2(x_2|t_2)}}\ . \qquad (5.2.9)$$

By substituting (5.1.4) and (5.2.6) into (5.2.9) one obtains:

$$\rho(x_1,x_2|t_1,t_2) = \frac{t_1 t_2\mu_2(\lambda)}{\sqrt{t_1^2\mu_2(\lambda)+t_1\mu_1'(\lambda)}\ \sqrt{t_2^2\mu_2(\lambda)+t_2\mu_1'(\lambda)}} \qquad (5.2.10)$$

The substitution of the coefficient of dispersion for λ into (5.2.10) leads to (5.2.8).

Lemma 5.5

For the symmetrical case where $t_1 = t_2 = t$ we have

$$\rho(x_1,x_2|t_1,t_2) = 1 - \frac{1}{1+tCOD(\lambda)} . \tag{5.2.11}$$

The proof is trivial.

Lemma 5.6

For the observational period $t = 1$ the correlation coefficient can be expressed in terms of the coefficient of dispersion for the mixing distribution as follows:

$$\rho(x_1,x_2|1,1) = \frac{COD(\lambda)}{1+COD(\lambda)} . \tag{5.2.12}$$

The proof is obtained from Lemma 5.5.

Lemma 5.7

For the symmetrical case $t_1 = t_2 = t$ the following relation exists:

$$\rho(x_1,x_2|t,t) = \frac{t\ \rho(x_1,x_2|1,1)}{1 + (t-1)\ \rho(x_1,x_2|1,1)} . \tag{5.2.13}$$

Proof

From (5.1.8) one obtains:

$$\mathrm{COD}(\lambda) = \frac{\mathrm{COD}(x_1|t)-1}{t} = \frac{\mathrm{COD}(x_1|1)-1}{t} \tag{5.2.14}$$

and

$$\rho(x_1,x_2|t,t) = 1 - \frac{1}{\mathrm{COD}(x_1|t)} . \tag{5.2.15}$$

Substitution of (5.2.14) into (5.2.15) leads to:

$$\rho(x_1,x_2|t,t) = 1 - \frac{1}{1 + t[\mathrm{COD}(x_1|1)-1]} . \tag{5.2.16}$$

From (5.2.15) one obtains the coefficient of correlation for the observational period

$$\rho(x_1,x_2|1,1) = 1 - \frac{1}{\mathrm{COD}(x_1|t)} . \tag{5.2.17}$$

Substitution of (5.2.17) into (5.2.16) leads to (5.2.13).

<u>Lemma 5.8</u>

The regression curve $E(x_2|x_1)$ for bivariate Poisson mixtures is a linear regression if and only if $\Psi(\lambda)$ is the Gamma distribution with parameters γ and θ.

Proof

The conditional distribution of $(x_2|x_1,t_1,t_2)$ is given by:

$$f(x_2|x_1,t_1,t_2) = \frac{f(x_1,x_2|t_1,t_2)}{f(x_1|t_1)} = \tag{5.2.18}$$

$$= \frac{\displaystyle\int_0^\infty e^{-(t_1+t_2)\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \frac{(t_2\lambda)^{x_2}}{x_2!} \Psi(\lambda)\, d\lambda}{\displaystyle\int_0^\infty e^{-t_1\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \Psi(\lambda)\, d\lambda} .$$

Consequently the regression curve is:

$$E(x_2|x_1,t_1,t_2) = \frac{\displaystyle\sum_{x_2=0}^{\infty} x_2 \int_0^\infty e^{-(t_1+t_2)\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \frac{(t_2\lambda)^{x_2}}{x_2!} \Psi(\lambda)\, d\lambda}{\displaystyle\int_0^\infty e^{-t_1\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \Psi(\lambda)\, d\lambda}$$

$$= \frac{t_2}{t_1} (x_1+1) \frac{\displaystyle\int_0^\infty e^{-t_1\lambda} \frac{(t_1\lambda)^{x_1+1}}{(x_1+1)!} \Psi(\lambda)\, d\lambda}{\displaystyle\int_0^\infty e^{-t_1\lambda} \frac{(t_1\lambda)^{x_1}}{x_1!} \Psi(\lambda)\, d\lambda}$$

$$= \frac{t_2}{t_1} (x_1+1) \frac{f(x_1+1|t_1)}{f(x_1|t_1)} . \tag{5.2.19}$$

Let us denote the ratio function of univariate compound Poisson as:

$$R(x_1|t_1) = \frac{f(x_1+1|t_1)}{f(x_1|t_1)} . \tag{5.2.20}$$

If $\Psi(\lambda)$ is the Gamma distribution with parameters γ and θ then $f(x_1|t_1)$ is the NBD model with parameters γ and θ_{t_1}, where

$$\theta_{t_1} = \frac{t_1\theta}{1 + (t_1-1)\theta} . \tag{5.2.21}$$

The ratio function of NBD with parameters γ and θ_{t_1} is:

$$R(x_1|t_1) = \frac{x_1+\gamma}{x_1+1}\,\theta_{t_1} \tag{5.2.22}$$

Substitution of (5.2.22) into (5.2.19) leads to:

$$E(x_2|x_1,t_1,t_2) = \frac{t_2}{t_1}\,\theta_{t_1}\,(x_1+\gamma) \tag{5.2.23}$$

which is a linear regression. It should be noted also that

$$E(x_2|x_1=0,t_1,t_2) = \frac{t_2}{t_1}\,\gamma\,\theta_{t_1} . \tag{5.2.24}$$

The purpose of the second part of the proof is to show that if $\Psi(\lambda)$ is not a Gamma type probability density function then the regression of x_2 on x_1 is not linear (Johnson, 1957).

Equation (5.2.19) can be written also as

$$E(x_1|x_2,t_1,t_2) = \frac{t_2}{t_1}\,\frac{\int_0^\infty e^{-t_1\lambda}\,(t_1\lambda)^{x_1+1}\,\Psi(\lambda)\,d\lambda}{\int_0^\infty e^{-t_1\lambda}\,(t_1\lambda)^{x_1}\,\Psi(\lambda)\,d\lambda} . \tag{5.2.25}$$

For the regression x_2 on x_1 to be linear we must have

$$a + bx_1 = \frac{t_2}{t_1}\,\frac{\int_0^\infty e^{-t_1\lambda}\,(t_1\lambda)^{x_1+1}\,\Psi(\lambda)\,d\lambda}{\int_0^\infty e^{-t_1\lambda}\,(t_1\lambda)^{x_1}\,\Psi(\lambda)\,d\lambda} , \quad \text{or} \tag{5.2.26}$$

$$\alpha + \beta x_1 = \frac{\int_0^\infty e^{-t_1\lambda}\,(t_1\lambda)^{x_1+1}\,\Psi(\lambda)\,d\lambda}{\int_0^\infty e^{-t_1\lambda}\,(t_1\lambda)^{x_1}\,\Psi(\lambda)\,d\lambda} . \tag{5.2.27}$$

Equation (5.2.27) can be written as

$$\alpha + \beta x_1 = \frac{\mu'_{x_1+1}(\lambda|t_1)}{\mu'_{x_1}(\lambda|t_1)}, \quad \text{where} \tag{5.2.28}$$

μ'_{x_1} and μ'_{x_1+1} are respectively the x_1th and (x_1+1)th moments about zero of the probability density function

$$\frac{e^{-t_1\lambda}\,\varphi(\lambda)}{\int_0^\infty e^{-t_1\lambda}\,\varphi(\lambda)\,d\lambda}, \quad \text{for } \lambda > 0\,. \tag{5.2.29}$$

Since $\mu'_0(\lambda|t_1) = 1$ it follows that

$\mu'_1(\lambda|t_1) = \alpha$, $\mu'_2(\lambda|t_1) = \alpha(\alpha+\beta)$, $\mu'_3(\lambda|t_1) = \alpha(\alpha+\beta)(\alpha+2\beta)$

and in general $\mu'_i = \alpha(\alpha+\beta)\ldots[\alpha+(i-1)\beta]$.

Let us denote $\beta' = \alpha + \beta$ and $\beta'_i = \alpha + (i-1)\beta$. Hence we have

$\mu'_i < (\beta'_i)^i$ and $\lim\limits_{i\to\infty} \left[\dfrac{\mu_i}{i}\right]^{1/i} < \beta'$. Hence distribution (5.2.29) is uniquely determined by the moments $\{\mu'_i\}$ and $\varphi(\lambda)$ is uniquely detemined by (5.2.26). A function $\varphi(\lambda)$ of Gamma type does satisfy (5.2.26) and is therefore the only probability density function which leads to a linear regression of x_2 on x_1.

Example

We shall consider a compound Poisson distribution with the GIG as the mixing distribution

$$\varphi(\lambda) = \frac{\left[\dfrac{2\sqrt{1-\theta}}{\alpha\theta}\right]^{\gamma}}{2K_\gamma(\alpha\sqrt{1-\theta})}\,\lambda^{\gamma-1}\,e^{-\left[\frac{1-\theta}{\theta}\right]\lambda - \frac{\alpha^2\theta}{4\lambda}}, \quad \text{where} \tag{5.2.28}$$

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

The resulting bivariate Poisson mixture is

$$f(x_1,x_2|t_1,t_2) = \frac{1}{K_\gamma(\alpha\sqrt{1-\theta})}\left[\frac{1-\theta}{1+(t_1+t_2-1)\theta}\right]^{\gamma/2}\frac{t_1^{x_1}\,t_2^{x_2}}{x_1!\,x_2!} \qquad (5.2.29)$$

$$\cdot\left[\frac{\alpha\theta}{2\sqrt{1+(t_1+t_2-1)\theta}}\right]^{x_1+x_2} K_{x_1+x_2+\gamma}(\alpha\sqrt{1+(t_1+t_2-1)\theta}),$$

where $x_1 \ge 0$ and $x_2 \ge 0$.

The marginal distributions of x_1 and x_2 are univariate Generalized Inverse Gaussian-Poisson distributions. The marginal distribution of x_1 is therefore:

$$f(x_1|t_1) = \frac{1}{K_\gamma(\alpha\sqrt{1-\theta})}\left[\frac{1-\theta}{1+(t_1-1)\theta}\right]^{\gamma/2}\frac{1}{x_1!} \qquad (5.2.30)$$

$$\cdot\left[\frac{t_1\alpha\theta}{2\sqrt{1+(t_1-1)\theta}}\right]^{x_1} K_{x_1+\gamma}(\alpha\sqrt{1+(t_1-1)\theta}).$$

The conditional probability density function of x_2 on x_1 is:

$$f(x_2|x_1,t_1,t_2) = \frac{1}{K_{x_1+\gamma}(\alpha\sqrt{1+(t_1-1)\theta})}\left[\frac{1+(t_1-1)\theta}{1+(t_1+t_2-1)\theta}\right]^{\frac{x_1+\gamma}{2}} \qquad (5.2.31)$$

$$\cdot\frac{1}{x_2!}\left[\frac{t_2\alpha\theta}{2\sqrt{1+(t_1+t_2-1)\theta}}\right]^{x_2} K_{x_1+x_2+\gamma}(\alpha\sqrt{1+(t_1+t_2-1)\theta}).$$

From (5.2.31) we find the regression curve of x_2 on x_1:

$$E(x_2|x_1,t_1,t_2) = \frac{t_2\alpha\theta}{2\sqrt{1+(t_1-1)\theta}}\,\frac{K_{x_1+\gamma+1}(\alpha\sqrt{1+(t_1-1)\theta})}{K_{x_1+\gamma}(\alpha\sqrt{1+(t_1-1)\theta})}, \qquad (5.2.32)$$

which is a non-linear regression.

Lemma 5.9

The ascending diagonal arrays for the bivariate Poisson mixtures are distributed following the Binomial distribution.

Proof

The bivariate Poisson mixtures are defined as:

$$f(x_1,x_2|t_1,t_2) = \int_0^\infty e^{-(t_1+t_2)\lambda} \frac{(t_1\lambda)^{x_1}(t_2\lambda)^{x_2}}{x_1!x_2!} \varphi(\lambda)\, d\lambda$$

$$= \frac{t_1^{x_1} t_2^{x_2}}{x_1!x_2!} \int_0^\infty e^{-(t_1+t_2)\lambda} \lambda^{x_1+x_2} \varphi(\lambda)\, d\lambda$$

$$= \frac{t_1^{x_1} t_2^{x_2}}{x_1!x_2!} f(S|t_1.t_2) \text{, where} \qquad (5.2.33)$$

$$f(S|t_1,t_2) = \int_0^\infty e^{-(t_1+t_2)\lambda} \lambda^S \varphi(\lambda)\, d\lambda \quad \text{and} \qquad (5.2.34)$$

$S = x_1+x_2$.

The probability density function for the ascending diagonal arrays is defined as:

$$\phi(x_1|S,t_1,t_2) = \frac{f(x_1|S,t_1,t_2)}{\sum_{x_1=0}^{S} f(x_1|S,t_1,t_2)}$$

$$= \frac{\frac{t_1^{x_1} t_2^{S-x_1}}{x_1!(S-x_1)!} f(S|t_1,t_2)}{\sum_{x_1=0}^{S} \frac{t_1^{x_1} t_2^{S-x_1}}{x_1!(S-x_1)!} f(S|t_1,t_2)} =$$

$$= \frac{\binom{S}{x_1} t_1^{x_1} t_2^{S-x_1}}{(t_1+t_2)^S}$$

$$= \binom{S}{x_1} \left[\frac{t_1}{t_1+t_2}\right]^{x_1} \left[1 - \frac{t_1}{t_1+t_2}\right]^{S-x_1} . \qquad (5.2.35)$$

Since the last expression is the probability density function of the Binomial distribution with the parameter $p = \frac{t_1}{t_1+t_2}$, it follows that for the symmetrical case when $t_1 = t_2$ the distribution for the ascending diagonal arrays will be Binomial with the parameter $p = 1/2$.