A COM-Poisson type generalization of the Binomial distribution and its properties and applications

Patrick Borges[a], Josemar Rodrigues[b] and Narayanaswamy Balakrishnan[c]

[a]Department of Statistics, Universidade Federal do Espírito Santo, Vitøria, Brazil

[b]Department of Mathematics and Statistics, Universidade de São Paulo, São Paulo, Brazil

[c]Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

Abstract

Shmueli *et al.* (2005) introduced a new discrete distribution, called COM-Poisson-Binomial distribution, by adding a dispersion parameter to the Binomial distribution. However, they did not study the mathematical properties of their proposed family of distributions. In this paper, we investigate for this probability distribution the moments, probability and moment generating functions and how the dispersion parameter affects the asymptotical approximation of the COM-Poisson-Binomial distribution to the COM-Poisson distribution. Inferential problems with two data sets are also considered for illustrative purposes.

Key words: COM-Poisson-Binomial distribution, dependent Bernoulli variables, Correlation coefficient, exponential family, Weighted Poisson distributions.

1. Introduction

Usually the binomial and Poisson distributions are used to analyze discrete data. However, it seems wise to consider flexible alternative models to take into account the overdispersion or underdispersion (see Hinde & Demetrio (1998)). Thus, the binomial and Poisson distributions have been generalized in several ways to handle the problem of dispersion inherent in the analysis of discrete data that may arise with the presence of aggregation of the individuals. For instance:

(i) in plant selection study the association among two plants arises when competing about the quantity of nutrients;

(ii) in biological study (see Yakovlev & Tsodikov (1996) and Borges *et al.* (2012)), it is usually assumed that cells in a tissue are independent. However, the biological independence assumption may not be true when the dynamics of the cell population of a normal tissue is considered. It is therefore desirable to construct new models with strong biological interpretation of the dependence incorporated in the carcinogenesis process.

*Corresponding author: Patrick Borges. Departamento de Estatística, Universidade Federal do Espírito Santo - UFES. Av. Fernando Ferrari 514, Goiabeiras, CEP 29075-910, Vitøria ES, Brazil. Email: patrickborges@yahoo.com.br

The binomial distribution has been generalized in various ways. Rudolfer (1990), Madsen (1993) and Luceño & Ceballos (1995) have summarized most of these generalizations. Among these extensions, there are the multiplicative and the additive generalized binomial distributions which were derived by Altham (1978). The probability mass function (pmf) of the multiplicative binomial distribution is a multiplication of its pmf by a factor. It makes the variance greater or less than the corresponding binomial variance depending on the values of the factor. On the other hand, the additive binomial distribution is a mixture of three conventional binomial models. Altham (1978) developed the correlated Binomial model by correcting the Binomial model via the method suggested by Bahadur (1961) to encompass dependent Bernoulli variables. A three-parameter binomial distribution was derived by Paul (1985, 1987), which is a generalization of the Binomial, beta-binomial and the correlated binomial distribution proposed by Kupper & Haseman (1978). Ng (1989) developed the modified binomial distributions. In this approach, the binomial distribution is changed and the resulting distribution becomes more spread out (indicating positive correlation among the Bernoulli variables), or more peaked (indicating negative correlation among the Bernoulli variables) than the binomial distribution. A four-parameter binomial distribution was derived by Fu & Sproule (1995). This new distribution assumes values between α and β for $\alpha < \beta$, rather than the usual values 0 or 1. Lindsey (1995) and Luceño & Ceballos (1995) proposed a generalized binomial distribution which is discussed in details in Diniz *et al.* (2010). Chang & Zelterman (2002) generalizes the binomial distribution by considering Bernoulli variables with probability of success depending on the previous one. Tsai *et al.* (2003) presented a model that studies the overall error rate when testing multiple hypotheses. This model involves the distribution of the sum of dependent Bernoulli trials, and it is approximated thorough a beta-binomial structure. Instead of using the beta-binomial model, Gupta & Tao (2010) derived the exact distribution of the sum of dependent Bernoulli variables and not identically distributed. Minkova & Omey (2011) defined a new binomial distribution related to the interrupted Markov chain. Another extension of the binomial distribution is the COM-Poisson-binomial distribution (CMPB, for short) introduced in Shmueli *et al.* (2005), however, they did not study the mathematical properties of their family which are studied in details in this paper. A recent application of CMPB distribution can be found in Kadane & Naeshagen (2013).

The CMPB distribution arises as the conditional distribution of a COM-Poisson variable (Conway & Maxwell, 1962) given a sum of two COM-Poisson variables with the same dispersion parameter. It generalizes the binomial distribution and can be interpreted as the sum of dependent Bernoulli variables with a specific joint distribution (see Remark 1). The dispersion parameter governs the correlation among the Bernoulli variables, overdispersion and underdispersion relative to binomial distribution. The CMPB distribution is appealing from a theoretical point of view since it belongs to the exponential and weighted Poisson families (Castillo & Pérez-Casany, 1998, 2005), and the sufficient statistic is defined by the mean and the log-geometric mean of the data. We refer to Barndorff-Nielsen (1978) for a general theory of exponential families.

The rest of this paper is organized as follows. In Section 2, we present the CMPB distribution with its mathematical properties. Section 3 describes the maximum likelihood method for estimating the parameters. In Section 4, we apply the CMPB distribution to two real datasets and show that this model provides an excellent fit to both these datasets. Finally, some concluding remarks are made in Section 5.

2. The CMPB distribution and its properties

The probability mass function (pmf) of the CMPB distribution (Shmueli *et al.*, 2005) is given by

$$\mathbb{P}[X = x|m, p, \nu] = \frac{\binom{m}{x}^{\nu} p^x (1-p)^{m-x}}{\sum_{k=0}^{m} \binom{m}{k}^{\nu} p^k (1-p)^{m-k}}, \quad x = 0, 1, \ldots, m, \tag{1}$$

for $m \in \mathbb{Z}^+$ (set of known non-negative integers), $p \in (0, 1)$ and $\nu \in \Re$. For $\nu = 1$ we have the binomial distribution. The values of $\nu > 1$ correspond to underdispersion, whereas the values of $\nu < 1$ represent overdispersion with respect to the binomial distribution. For $\nu \to \infty$, the pmf is concentrated at the point mp and for $\nu \to -\infty$ is concentrated at 0 or m. Figure 2.1 presents the pmf of the CMPB distribution for $m = 6$ and different values of p and ν.

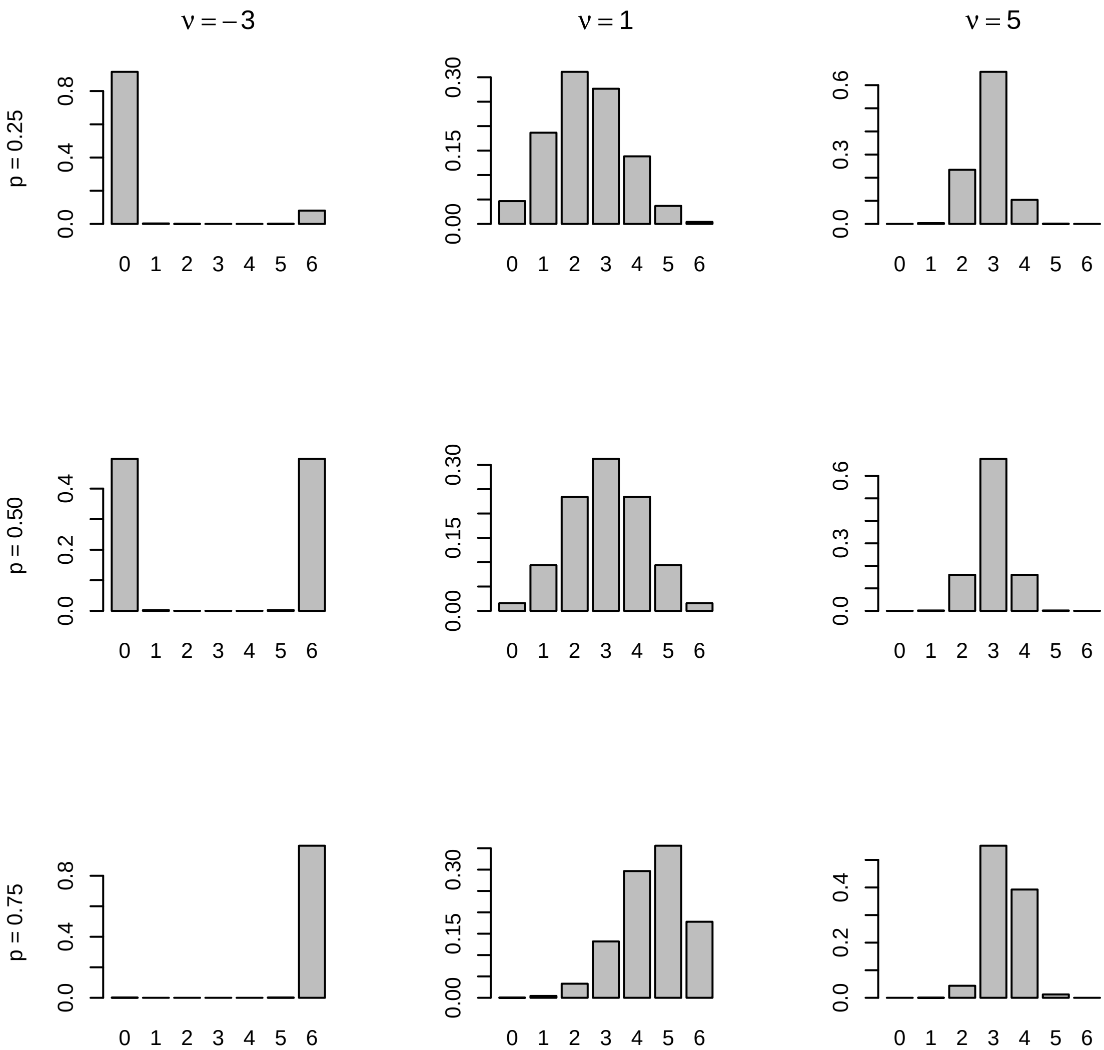

Figure 2.1: COM-Poisson-binomial for $m = 6$ and various choices of p and ν.

Remark 1. The CMPB distribution can be interpreted as a sum of equicorrelated Bernoulli variables Z_i $(i = 1, \ldots, m)$ with joint distribution

$$\mathbb{P}[Z_1 = z_1, \ldots, Z_m = z_m] = \frac{\binom{m}{x}^{\nu-1} p^x (1-p)^{m-x}}{\sum_{z_1=0}^{1} \cdots \sum_{z_m=0}^{1} \binom{m}{x}^{\nu-1} p^x (1-p)^{m-x}}, \quad \boldsymbol{z} = (z_1, \ldots, z_m) \in \{0,1\}^m, \quad (2)$$

where $x = \sum_{i=1}^{m} z_i$. The measure of linear association between the Bernoulii variables is given by

$$\rho = \frac{p(1-p)(1-4^{\nu-1})}{(p + (1-p)2^{\nu-1})(1 - p(1-2^{\nu-1}))}, \quad i \neq j, \quad i, j = 1, \ldots, m. \quad (3)$$

Thus, for $\nu > 1$, the Bernoulli variables are in different directions, and for $\nu < 1$ in the same direction.

Remark 2. The COM-Poisson distribution (Conway & Maxwell, 1962) is an approximation to the CMPB distribution when m is getting large and

$$\lambda = m^{\nu} p, \tag{4}$$

for $\nu \geq 0$. In other words, for $m \to \infty$ we have that

$$\begin{aligned} \lim_{m\to\infty} \mathbb{P}[X = x|m, p, \nu] &= \lim_{m\to\infty} \frac{\binom{m}{x}^{\nu} \left(\frac{\lambda}{m}\right)^{x} \left(1 - \frac{\lambda}{m}\right)^{m-x}}{\sum_{k=0}^{m} \binom{m}{k}^{\nu} \left(\frac{\lambda}{m}\right)^{k} \left(1 - \frac{\lambda}{m}\right)^{m-k}} \\ &= \frac{\lambda^{x}}{(x!)^{\nu}} \frac{1}{\sum_{j=0}^{\infty} \frac{\lambda^{j}}{(j!)^{\nu}}}, \quad x = 0, 1, \ldots, \end{aligned} \tag{5}$$

which is the pmf of the COM-Poisson distribution. This result is a generalization of the Binomial approximation to the Poisson distribution where the dispersion parameter ν in (4) indicates the influence of the superdispersion or underdispersion on the approximation to the COM-Poisson distribution. This is an interesting result not mentioned by Shmueli *et al.* (2005) and one of the main contribution of this paper.

In order to obtain a COM-Poisson type structure for the pmf of the CMPB distribution, we divide the numerator and denominator of (1) by a factor $(1-p)^{m}(m!)^{\nu}$, and let us consider the following parametrization:

$$\theta = \frac{p}{1-p}. \tag{6}$$

Under the parametrization (6) the pmf of the CMPB distribution, denoted by $X \sim$CMPB(m,θ,ν), is given by:

$$\mathbb{P}[X = x|m, \theta, \nu] = \frac{1}{Z(\theta, \nu)} \frac{\theta^{x}}{[x!(m-x)!]^{\nu}}, \quad x = 0, 1, \ldots, m, \tag{7}$$

where $Z(\theta, \nu) = \sum_{j=0}^{m} \frac{\theta^{j}}{[j!(m-j)!]^{\nu}}$. Also, the pmf (7) can be rewritten as

$$\mathbb{P}[X = x|m, \theta, \nu] = \frac{\exp\big(x \log(\theta) - \nu \log(x!(m-x)!)\big)}{Z(\theta, \nu)}. \tag{8}$$

It follows from (8) that the CMPB distribution belongs to the *full exponential family* on $\mathbb{Z}^{+}$, where $T(X) = \big(X, \log(X!(m-X)!)\big)$ is the sufficient statistic and $(\log\theta, \nu) \in \mathbb{R}^{2}$ its corresponding natural parameters (see Barndorff-Nielsen (1978)). Also, the CMPB distribution belongs to the family of weighted Poisson distributions (Castillo & Pérez-Casany, 1998, 2005) defined as follows:

$$\mathbb{P}[X = x|\theta, \nu] = \frac{w(x;\nu) p^{*}(x;\theta)}{\mathbb{E}_{\theta}[w(X;\nu)]}, \tag{9}$$

where $w(\cdot\ ;\ \nu)$ is a non-negative weight function with parameter ν, $p^{*}(\cdot\ ;\ \theta)$ is the pmf of a Poisson distribution with parameter θ, and $\mathbb{E}_{\theta}[\cdot]$ indicates that the expectation is taken with respect to the Poisson distribution with parameter θ. Therefore, if we take the weighted function as

$$w(x;\nu) = \begin{cases} \frac{[x!(m-x)!]^{1-\nu}}{(m-x)!} & , \ if \quad x \leq m \\ 0 & , \ if \quad x > m \end{cases},$$

the CMPB distribution (7) can be expressed as the weighted Poisson distribution in (9). So, some interesting characterizations of the overdispersion and underdispersion for the CMBP distributions can be derived from the weighted Poisson distributions.

Rodrigues *et al.* (2009) have shown that the probability generating function (pgf) of the weighted Poisson variable is given by

$$\begin{aligned} \mathbb{A}_X(s) &= \mathbb{E}[s^X] = \sum_{x=0}^{\infty} s^x \frac{w(x;\nu)p^*(x;\theta)}{\mathbb{E}_\theta[w(X;\nu)]} \\ &= \exp\big(-\theta(1-s)\big)\frac{\mathbb{E}_{\theta s}[w(X;\nu)]}{\mathbb{E}_\theta[w(X;\nu)]}, \quad \text{for} \quad 0 \le s \le 1. \end{aligned} \tag{10}$$

From (10), the pgf of the CMPB distribution is given by

$$\mathbb{A}_X(s) = \exp\big(-\theta(1-s)\big)\frac{\mathbb{E}_{\theta s}\left[\frac{[X!(m-X)!]^{1-\nu}}{(m-X)!}\right]}{\mathbb{E}_{\theta}\left[\frac{[X!(m-X)!]^{1-\nu}}{(m-X)!}\right]} \tag{11}$$

$$= \frac{Z(\theta s,\nu)}{Z(\theta,\nu)}, \quad \text{for} \quad 0 \le s \le 1. \tag{12}$$

3. Maximum likelihood estimation of the parameters

Let $\mathbf{X} = (X_1, \ldots, X_n)^\top$ be a random sample size of n with observed values $\mathbf{x} = (x_1, \ldots, x_n)^\top$ from a CMPB distribution with parameters θ and ν. Let us define

$$t_1 = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad t_2 = \frac{1}{n}\sum_{i=1}^{n} \log\left(x_i!(m-x_i)!\right).$$

Note that, t_1 and t_2 are, respctively, the sample mean and the log-geometric mean.

The log-likelihood function for the CMPB model based on the obseved sample $\mathbf{x}$ is

$$\ell(\theta,\nu;\mathbf{x}) = n\Big(\log(\theta)t_1 - \nu t_2 - \log\big(Z(\theta,\nu)\big)\Big). \tag{13}$$

The likelihood equations may be written as

$$\begin{cases} \theta \frac{d}{d\theta}\log\big(Z(\theta,\nu)\big) = \mathbb{E}[X] = t_1 \\ -\frac{d}{d\nu}\log\big(Z(\theta,\nu)\big) = \mathbb{E}\left[\log\big(x!(m-x)!\big)\right] = t_2 \end{cases}. \tag{14}$$

As mentioned in Section 2, the CMPB distribution is a member of the *full exponential family* on $\mathbb{Z}^+$, and $T(X) = \big(X, \log(X!(m-X)!)\big)$ is its corresponding sufficient statistic. Since $S = T(\mathbb{Z}^+)$ is not included in an affine subspace of $\mathbb{R}^2$, $\bar{T} = (T_1, T_2)$ is a minimal sufficient statistic for the CMPB distribution (see Barndorff-Nielsen (1978)). Let $\mathcal{T}$ be the interior of the convex hull of S. As the CMPB is a regular exponential distribution, it is well-known that there is a one-to one transformation from $\mathbb{R}^2$ to $\mathcal{T}$ expressed in terms of $\tau(\alpha,\nu) = \mathbb{E}_{(\alpha,\nu)}(T) = \big(\tau_1(\alpha,\nu), \tau_2(\alpha,\nu)\big)^\top$ (see Barndorff-Nielsen (1978)). So, if (t_1,t_2) belongs to $\mathcal{T}$, the maximum likelihood estimates (MLEs) of the parameters θ and ν are unique solution of the system equations (14). Since these equations cannot be solved analytically, an iterative method such as the Newton-Raphson method can be used (see Gelman *et al.* (1995), pages 272-273). In each iteration, the expectations, variances and covariances of $X!$ and $\log(X!(m-X)!)$ are computed by plugging the estimates for θ and ν obtained from the previous iterations in the expression

$$\mathbb{E}[h(X)] = \sum_{x=0}^{m} h(x) \frac{1}{Z(\theta,\nu)} \frac{\theta^x}{[x!(m-x)!]^{\nu}}. \tag{15}$$

The MLEs of the parameters can also obtained by direct maximization of the log-likelihood function (13) by using the SAS (PROC NLMIXED) or MaxBFGS routine of the Ox program (Doornik, 2013) or *optim* routine of the R package (R Development Core Team, 2013).

4. Fitting the COM-Poisson-binomial distribution

To illustrate the usefulness and flexibility of the CMPB distribution, we consider two real datasets. In the first example, the dataset consists of the number of male children in families taken from the hospital records in the nineteenth century Saxony (Sokal & Rohlf, 1994), and in the second example the secondary association of chromosomes in Brassika (Skellam, 1948). Furthermore, the binomial and beta-binomial (BB) distributions were also fitted in both examples by using the maximum likelihood method. The pmf of the BB distribution is given by

$$\mathbb{P}[X = x|m,\alpha,\beta] = \binom{m}{x} \frac{B(x+\alpha, n-x-\beta)}{B(\alpha,\beta)}, \quad x = 0,1,\dots,m,$$

where $\alpha > 0$, $\beta > 0$ and $B(\cdot,\cdot)$ denotes the beta function.

Example 1: The data in this example refer to the number of male children among the first 12 children of family size 13 in 6115 families taken from the hospital records in the nineteenth century Saxony (Sokal & Rohlf (1994), Lindsey (1995), p. 59). The thirteenth child is ignored to assuage the effect of families non-randomly stopping when a desired gender is reached. Thus, $m = 12$ is the family size and X is the number of male children. The expected frequencies, maximized kernel of the log-likelihood and estimates of the parameters under the binomial, BB and CMPB distributions are given in Table 4.1. The chi-square goodness-of-fit, the likelihood ratio (2Λ) and the kernel of the log-likelihood values are showed in Table 4.1 describing how well the CMPB and BB fit to the data.

Table 4.1: The goodness of fit of the binomial, BB and CMPB distributions

No. of males children		0	1	2	3	4	5	6	7	8	9	10	11	12	Kernel of the log-likelihood	Parameter Estimates	Chi-square Goodness-of-fit	values: 2Λ
Observed frequency		3	24	104	286	670	1033	1343	1112	829	478	181	45	7	-12485.67	—	—	—
	Binomial	0.9	12.1	71.8	258.5	628.1	1085.2	1367.3	1265.6	854.2	410.0	132.8	26.1	2.3	-12534.17	$\hat{p}$=0.5192	105.9498	—
Expected frequency	BB	2.3	22.6	104.8	310.9	655.7	1036.2	1257.9	1182.1	853.6	461.9	177.9	43.8	5.2	-12492.87	$\hat{\alpha}$=34.0350 $\hat{\beta}$=31.5160	13.9281	82.60
	CMPB	2.7	23.3	104.7	308.8	653.6	1037.6	1262.3	1184.0	850.8	458.6	177.5	45.0	5.9	-12492.35	$\hat{\theta}$=1.0682 $\hat{\nu}$=0.8433	13.1681	83.64

The kernel of the log-likelihood is obtained as follows: Let $P_x = \mathbb{P}[X = x]$ and O_x = observed frequency of $X = x$. Then, the kernel of the log-likelihood of any model is given by $\sum_{x=0}^{m} O_x \log(P_x) = \sum_{x=0}^{m} O_x \log\left(\frac{n\hat{P}_x}{n}\right) = \sum_{x=0}^{m} O_x \log E_x - n\log(n)$, where $n = \sum_{x=0}^{m} O_x$ (total frequency) and E_x is the expected frequency of $X = x$. Thus, it follows immediately that the kernel of the log-likelihood is $\sum_{x=0}^{m} O_x \log(O_x) - n\log(n)$.

The BB and the CMPB distributions give almost identical fits to the data (the CMPB distribution gives slightly better fit than the BB distribution). The likelihood ratio chi-square values of the BB and COMPB distributions are $-2(-12485.67 + 12492.87) = 82.60$ and $-2(-12485.67 + 12492.35) = 83.64$, respectively. These values show that both the BB and CMPB models give significant improvements over the binomial distribution (the best fit is evident especially among the tails).

Example 2: The data in this example refer to 337 observations on the secondary association of chromosomes in Brassika; $m = 3$ is the number of chromosomes and X is the number of pairs of bivalents showing association. Skellam fitted the BB distribution showing a remarkable improvement fit over the Binomial distribution. We fit the CMPB distribution to these data and obtained a similar result. The observed frequencies, the expected frequencies, the kernel of the log-likelihood and the estimates of the parameters are presented in Table 4.2.

Table 4.2: The goodness of fit of the binomial, BB and CMPB distributions for the Skellam's Brassika data

No. of associations	Observed frequency	Expected frequency		
		Binomial	BB	CMPB
0	32	24.86	33.43	33.99
1	103	103.25	97.56	97.02
2	122	142.94	128.53	127.98
3	80	65.96	77.47	78.00
Kernel of the log-likelihood	-436.44	-440.38	-436.82	-436.84
Parameter estimates	—	$\hat{p} = 0.5806$	$\hat{\alpha} = 6.5375$ $\hat{\beta} = 4.7213$	$\hat{\theta} = 1.3190$ $\hat{\nu} = 0.7026$
Chi-square values:				
Goodness-of-fit	—	8.10	0.78	0.82
2Λ	—	—	7.12	7.08

Looking at the expected frequencies and the maximized kernel, we see that both the BB and CMPB models give significant improvements over the Binomial model.

5. Concluding Remarks

We study and discuss here the mathematical properties of the CMPB distribution proposed by Shmueli *et al.* (2005) as a two-parameter extension of the binomial distribution. The main advantage of this model is its flexibility to handle overdispersion or underdispersion commonly encountered in count data sets. The CMPB distribution is appealing from a theoretical point of view since it belongs to the exponential family as well as to the weighted Poisson distributions family. Various statistical and probabilistic properties were derived such as moments, probability and moment generating functions. It is showed how the dispersion parameter affects the approximation of the CMPB distribution to the COM-Poisson distribution as the sample size gets large. Applications of the CMPB distribution thorough the maximum likelihood estimation to a real datasets show that the CMPB distribution can yield a better fit than some well-known models. Since the CMPB distribution belongs to exponential family we believe that is possible to develop a subjective or objective Bayesian analysis for this model. Work in this direction is currently under progress and we hope to report these findings in a future paper.

References

Altham, P. M. E. (1978). Two generalizations of the binomial distribution. *Journal of the Royal Statistical Society, Series C*, **27**, 162–167.

Bahadur, R. R. (1961). *A representation of the joint distribution of responses to n dichotomous items. In: Studies in Item Analysis and Prediction*. Solomon, California.

Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. Wiley, Chichester, UK.

Borges, P., Rodrigues, J. & Balakrishnan, N. (2012). Correlated destructive generalized power series cure rate models and associated inference with an application to a cutaneous melanoma data. *Computational Statistics & Data Analysis*, **56**, 1703–1713.

Castillo, J. & Pérez-Casany, M. (1998). Weighted Poisson distributions for overdispersion and underdispersion situations. *Annals of the Institute of Statistical Mathematics*, **50**, 567–585.

Castillo, J. & Pérez-Casany, M. (2005). Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference*, **134**, 486–500.

Chang, Y. & Zelterman, D. (2002). Sums of dependent Bernoulli random variables and disease clustering. *Statistics & Probability Letters*, **57**, 363–373.

Conway, R. W. & Maxwell, W. L. (1962). A queuing model with state dependent service rates. *Journal of Industrial Engineering*, **12**, 132–136.

Diniz, C. A. R., Tutia, M. H. & Leite, J. G. (2010). Bayesian analysis of a correlated binomial model. *Journal of Probability and Statistics*, **24**, 68–77.

Doornik, J. (2013). *Ox 5: object-oriented matrix programming language, 5th ed*. Timberlake Consultants, London.

Fu, J. & Sproule, R. (1995). A generalization of the binomial distribution. *Communications in Statistics - Theory and Methods*, **24**, 2645–2658.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman and Hall, New York.

Gupta, R. C. & Tao, H. (2010). A generalized correlated binomial distribution with application in multiple testing problems. *Metrika*, **71**, 59–77.

Hinde, J. & Demetrio, C. G. B. (1998). Overdispersion: Models and estimation. *Computational Statistics & Data Analysis*, **27**, 151–170.

Kadane, J. B. & Naeshagen, F. (2013). The number of killings in rural Norway, 1300-1569. *Annals of Applied Statistics*, **(to appear)**.

Kupper, L. L. & Haseman, J. K. (1978). The use of a correlated binomial model for the analysis of certain toxicological experiments. *Biometrics*, **34**, 69–76.

Lindsey, J. K. (1995). *Modelling Frequency and Count Data*. Oxford University Press, Oxford, UK.

Luceño, A. & Ceballos, F. (1995). Describing extra-binomial variation with partially correlated models. *Communications in Statistics - Theory and Methods*, pages 1637–1653.

Madsen, R. W. (1993). Generalized binomial distributions. *Communications in Statistics - Theory and Methods*, **22**, 3065–3086.

Minkova, L. D. & Omey, E. (2013). A new Markov binomial distribution. *Communications in Statistics - Theory and Methods*, **(to appear)**.

Ng, T.-H. (1989). A new class of modified binomial distributions with applications to certain toxicological experiments. *Communications in Statistics - Theory and Methods*, **18**, 3477–3492.

Paul, S. R. (1985). A three-parameter generalization of the binomial distribution. *Communications in Statistics - Theory and Methods*, **14**, 1497–1506.

Paul, S. R. (1987). On the beta-correlated binomial (bcb) distribution - a three-parameter generalization of the binomial distribution. *Communications in Statistics - Theory and Methods*, **16**, 1473–1478.

Rodrigues, J., de Castro, M., Cancho, V. & Balakrishnan, N. (2009). COM-Poisson cure rate survival models and an application to a cutaneous melanoma data. *Journal of Statistical Planning and Inference*, **139**, 3605–3611.

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rudolfer, S. M. (1990). A Markov chain model of extrabinomial variation. *Biometrika*, **7**, 255–264.

Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S. & Boatwright, P. (2005). A useful distribution for fitting discrete data: revival of the Conway-Maxwell-Poisson distribution. *Journal of the Royal Statistical Society, Series C*, **54**, 127–142.

Skellam, J. G. (1948). A probability distribution derived from the Binomial distribution by regarding the probability of sucess as variable between the sets of trials. *Journal of the Royal Statistical Society, Series B*, **10**, 257–261.

Sokal, R. R. & Rohlf, F. J. (1994). *Biometry: The Principles and Practices of Statistics in Biological Research*. W H Freeman & CO, New York.

Tsai, C., Hsueh, H. & Chen, J. J. (2003). Estimation of false discovery rates in multiple testing: application to genemicroarray data. *Biometrics*, **59**, 1071–1081.

Yakovlev, A. & Tsodikov, A. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific Publishers, Singapore.