Commun. Math. Res.
doi: 10.4208/cmr.2021-0045

Vol. **38**, No. 1, pp. 28-51
February 2022

The Binomial-Discrete Poisson-Lindley Model: Modeling and Applications to Count Regression

Christophe Chesneau[1], Hassan S. Bakouch[2], Yunus Akdoğan[3,*] and Kadir Karakaya[3]

[1] *LMNO, University of Caen-Normandie, Caen, France.*
[2] *Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt.*
[3] *Department of Statistics, Faculty of Science, Selcuk University, Konya, Turkey.*

Received 13 April 2021; Accepted 6 September 2021

Abstract. On the basis of a well-established binomial structure and the so-called Poisson-Lindley distribution, a new two-parameter discrete distribution is introduced. Its properties are studied from both the theoretical and practical sides. For the theory, we discuss the moments, survival and hazard rate functions, mode and quantile function. The statistical inference on the model parameters is investigated by the maximum likelihood, moments, proportions, least square, and weighted least square estimations. A simulation study is conducted to observe the performance of the bias and mean square error of the obtained estimates. Then, applications to two practical data sets are given. Finally, we construct a new flexible count data regression model called the binomial-Poisson Lindley regression model with two practical examples in the medical area.

AMS subject classifications: 60E05, 62E10, 62E15

Key words: Binomial-discrete family, Poisson-Lindley distribution, estimation, data analysis, count regression.

---

*Corresponding author. *Email address:* `yakdogan@selcuk.edu.tr` (Y. Akdoğan)

1 Introduction

In physics, counts are encountered as radioactive decay, or photon counting using a Geiger tube, and for modelling such physical issue, Hu *et al.* [9] introduced the binomial-discrete family of discrete distributions characterized by the compounding of two discrete distributions: The binomial distribution and a generic discrete distribution with support $\mathbb{N}=\{0,1,\ldots\}$. It is defined by the following probability mass function (PMF):

$$g(x;p,\zeta)=\sum_{n=x}^{\infty}\binom{n}{x}p^x(1-p)^{n-x}P_\zeta(N=n), \tag{1.1}$$

where $p\in(0,1)$ and N denotes a discrete random variable with support $\mathbb{N}$, which may depend on a parameter vector denoted by ζ. Hu *et al.* [9] studied the special member of the family defined with N following the Poisson random variable with parameter λ, and showed that $g(x;p,\lambda)$ corresponds to the PMF of a Poisson distribution with parameter λp. But the Poisson distribution is not always suitable for modeling and analysis of data counts, because its mean and variance are equal. Therefore, there is a need to introduce other distributions for counts. Recently, Kus *et al.* [13] investigated another special member of the family that arises with N following the discrete Lindley distribution. The binomial-discrete Lindley distribution has seen the light. Then, it was proved that it could provide better fits than the former discrete Lindley distribution for six different data sets. Several extensions of the binomial-discrete family can be found in Akdogan *et al.* [1] and Deniz [7].

In this paper, we investigate a new special member of the binomial-discrete family from a statistical point of view and, based on it, we set up a new regression model. This member assumes that N follows a well-known extension of the discrete Lindley distribution: the so-called the Poisson-Lindley (PL) distribution. The motivations behind the PL distribution are recalled below. First of all, it is defined by the following PMF:

$$f(x;\theta)=\frac{\theta^2(x+\theta+2)}{(1+\theta)^{x+3}}, \quad x\in\mathbb{N}, \tag{1.2}$$

where $\theta>0$. The PL distribution possesses tractable probability functions, as well as desirable properties, such as unimodality, overdispersion and increasing hazard rate function (HRF). Thanks to its skewness and kurtosis features, it can provide a better alternative to the Poisson, geometric and negative binomial distributions for modelling purposes. Further theoretical facts and applications for the PL distribution can be found in Sankaran [17] and Shanker and Fesshaye [18].