A NEW COUNT DATA PROBABILITY MODEL: PROPERTIES AND APPLICATIONS

Morifat[1], Bilal Ahmad Para[2,*]

•

[1,2]Department of Mathematical Sciences, Islamic University of Science and Technology, Kashmir, India

[1]morifat13@gmail.com, [2,*]Corresponding author: parabilal@gmail.com

Abstract

In the field of count data analysis, over-dispersion poses significant challenges, often limiting the effectiveness of traditional Poisson model. To address this limitation, we propose a novel two parameter distribution as an extension of Poisson distribution namely two-Parameter Poisson Garima Distribution (TPPGD). This distribution enhances modeling flexibility for over-dispersed data, offering a superior fit for real-world datasets. In this paper we derive the theoretical properties of TPPGD, including its probability mass function, cumulative distribution function and different statistical properties. Parameter estimation has been done using the maximum likelihood method and the moment method. Finally, the validity of the proposed model is checked using different real world data sets.

Keywords: Poisson distribution, Two Parameter Garima distribution, Compounding, Count data.

1. Introduction

Numerous probability models are acquired by researchers in order to analyze a wide range of data from different industries, including engineering, agriculture, transportation, medical, and other sectors. Lots of well-known techniques are employed to serve the purpose of constructing new probability distributions. Discretization, the T-X family, and compounding are a few well-known methods that offer a very effective means of extending popular parametric families of distributions to match data sets that are not well-fitted by classical distributions. In statistics, compounding technique plays an important role in discussing complex phenomena by integrating multiple distributions to model real world data effectively. The compounding technique originated in the early 20th century and the main figures in advancing use of compounding distributions were Greenwood & Udny Yule [5] who introduced concepts like negative binomial distribution in the 20th century. The negative binomial arises as a compounding distribution when a Poisson variable is mixed with a gamma-distributed parameter to model over-dispersion in count data. Compounding involves combining two or more probability distributions, where one distribution is considered conditional on the parameters of another. If the mean of Poisson follows inverse Gaussian, the resulting distribution is Poisson inverse Gaussian (Holla [6]). Despite being extensively utilized across all domains, these models are not always able to effectively characterize the relationships between variables and cannot handle all forms of count data. In this context, new models are created to provide better outcomes and serve as substitutes for traditional count models.