Eur. J. Stat. 6 (2026) 5
doi: 10.28924/ada/stat.6.5

On a Discrete New Generalized Pareto-Based Regression Model for Count Data

Jiji Jose[1], K. Jayakumar[2,*]

[1] *Department of Statistics University of Calicut, Kerala, 673635, India*
jijijose2793@gmail.com

[2] *Department of Statistics, Cochin University of Science and Technology, Kerala, 682022, India*
jkumarguruvayoor@gmail.com

[*] *Correspondence: jkumarguruvayoor@gmail.com*

Abstract. Count data often exhibit overdispersion, heavy tails, and decreasing failure rates, which limit the applicability of classical Poisson regression models. In this paper, we develop a regression model based on the discrete new generalized Pareto (DNGP) distribution to better capture these features. The proposed model incorporates covariate effects through suitable link function and parameters are estimated using maximum likelihood estimation methods. Simulation studies, model comparison and real data applications demonstrate that the DNGP regression model provides a flexible and effective alternative for analyzing complex count data.

1. Introduction

Regression models for count data have considerable attention across diverse disciplines, including economics, biostatistics, actuarial science, demography, and the social sciences, owing to their wide range of practical applications. One of the earliest contributions in this area was made by [8], who introduced multiple linear regression analysis for count responses assumed to follow a Poisson distribution. Subsequently, [3] examined the use of the Poisson distribution within a nonlinear regression framework for count data, particularly in situations where the sample mean and variance are approximately equal. A major advancement in the modeling of count data was achieved through the development of generalized linear models by [13], which was later comprehensively elaborated by [12].

Several discrete Pareto-type models have been developed to accommodate count data exhibiting decreasing failure rates and heavy-tailed characteristics; however, most of these developments have primarily focused on distributional properties and data fitting rather than on regression modeling. Notable contributions include the discrete Pareto (DP) by [11], generalized discrete Pareto distribution studied by [1], the discrete generalized Pareto model

Received: 15 Jan 2026.
Key words and phrases. Count data regression model; Discrete Pareto distribution; Discrete new generalized Pareto distribution; Maximum likelihood estimation; Vuong test.

proposed by [15] for analyzing road crash counts at accident-prone locations, and the discrete three-parameter Burr type XII and discrete Lomax distributions introduced by [14] for modeling kidney cyst count data. Further extensions such as the discrete Pareto type IV distribution ([4]), the discrete new generalized Pareto (DNGP) distribution developed by [5], the discrete truncated Pareto model applied by [10], and another generalizations by [9], [6], have demonstrated the flexibility of discrete Pareto models in handling overdispersion and heavy tails. Despite their practical relevance, regression frameworks incorporating these discrete Pareto distributions remain relatively underexplored. Motivated by this gap, the present paper constructs a regression model based on the DNGP distribution, enabling the systematic incorporation of covariate information while retaining the distribution's ability to capture complex dispersion patterns commonly observed in real-world count data.

The reasons behind proposing this model:

- To establish generalized family for all existing DP models.
- To expand generalized linear model (GLM) frameworks to accommodate heavy-tailed count phenomena.
- To generalize, and extend existing discrete probabilistic models for complex datasets.

This paper is organized as follows. Section 2 introduces the DNGP model and explains some of its properties in Section 3. Section 4 discusses characterization of the DNGP model. In Section 5, introduces DNGP regression model, including parameter estimation via MLE and simulation study in Section 6 and an application to real data analysis illustrated in Section 7. Finally, Section 8 presents the concluding remarks and outlines the major limitations of the proposed model.

2. DNGP model

To discuss the development of DNGP distribution, it is essential to know a generalization of the Pareto distribution namely new generalized Pareto (NGP) distribution is introduced by [7]. The PDF of the NGP distribution is provided in equation (1).

$$g_X(x) = \frac{\alpha\beta^{\alpha}\theta(1-\gamma)\gamma^{\theta}}{1-\gamma^{\theta}} \frac{x^{\alpha\theta}-1}{(\gamma x^{\alpha}+(1-\gamma)\beta^{\alpha})^{\theta+1}}; x > \beta, \alpha, \beta, \gamma, \theta > 0. \tag{1}$$

The survival and failure rate functions of the NGP distribution are, respectively, given by

$$S_X(x) = \frac{1}{1-\gamma^{\theta}} - \frac{\gamma^{\theta}}{1-\gamma^{\theta}} \left[\frac{x^{\alpha\theta}}{[\gamma x^{\alpha}+(1-\gamma)\beta^{\alpha}]^{\theta}}\right], \quad x \geq \beta, \tag{2}$$

and

$$r_X(x) = \frac{\left(\frac{1-\gamma}{\gamma}\right)\frac{\theta\alpha\beta^{\alpha}}{x^{\alpha+1}}}{\left[\left(\frac{1-\gamma}{\gamma}\right)\left(\frac{\beta}{x}\right)^{\alpha}+1\right]\left\{\left[\left(\frac{1-\gamma}{\gamma}\right)\left(\frac{\beta}{x}\right)^{\alpha}+1\right]^{\theta}-1\right\}}, \quad x > \beta. \tag{3}$$

Although NGP distribution provides a highly flexible framework for modelling heavy-tailed data, its structure can be further extended to capture a wider range of tail behaviours and distributional shapes. Motivated by the flexibility and practical relevance of the NGP distribution, we develop a new discrete generalization through survival discretization, called the DNGP distribution.

Using the survival function of the NGP distribution, the PMF of the DNGP distribution can be obatined as

$$\begin{aligned} f_Z(z) =& P(Z=z) = S_X(z) - S_X(z+1) \\ =& \frac{\gamma^\theta}{1-\gamma^\theta}\left[\frac{(z+1)^{\alpha\theta}}{(\gamma(z+1)^\alpha + (1-\gamma)\beta^\alpha)^\theta} - \frac{z^{\alpha\theta}}{(\gamma z^\alpha + (1-\gamma)\beta^\alpha)^\theta}\right], \end{aligned}$$

where $z = [\beta], [\beta+1], [\beta+2], ...; \alpha, \beta, \theta > 0, 0 < \gamma < 1$ and is denoted by DNGP$(\alpha, \beta, \gamma, \theta)$.

Definition 1. *A non negative integer-valued random variable Z is said to follow DNGP distribution if its PMF $f_Z(z) = P(Z = z)$, is of the following*

$$f_Z(z) = \frac{1}{1-\gamma^\theta}\left(g(z+1)^{-\theta} - g(z)^{-\theta}\right) \tag{4}$$

where $g(z) = \frac{z^\alpha + (\frac{1-\gamma}{\gamma})\beta^\alpha}{z^\alpha}$, $\alpha > 0, \beta > 0, 0 < \gamma < 1, \theta > 0$ *and* $z = [\beta], [\beta+1], [\beta+2], ..$

Result 1. *If $\beta = 1, \theta = 1$ and $\gamma \to 1$, DNGP reduces to DP distribution.*

Proof. When $\beta = \theta = 1$ and $\gamma \to 1$, the PMF becomes ,

$$f_Z(z) = \left(\frac{1}{z}\right)^\alpha - \left(\frac{1}{z+1}\right)^\alpha,$$

which is the PMF of DP distribution. □

Figure 1 presents the PMF plots of DNGP distribution for different combinations of parameters α, β, γ and θ. The plots illustrate that increasing θ or decreasing γ tends to make the distribution more peaked or unimodal, while larger values of γ or smaller α produce heavier right tails and monotonically decreasing shapes. Overall, these plots highlights the flexibility of the DNGP model in capturing a wide range of discrete data behaviours- from symmetric and bell-shaped forms to skewed and long-tailed patterns.

Result 2. *The CDF $F_Z(z)$ of the DNGP distribution having PMF in equation (4) is the following*

$$F_Z(z) = 1 - \frac{1}{1-\gamma^\theta}\left(1 - g(z+1)^{-\theta}\right). \tag{5}$$

Remark 1.

- $F_Z(0) = 1 - \frac{1}{1-\gamma^\theta}\left(1 - g(0)^{-\theta}\right)$.
- *The proportion of positive values:* $1 - F_Z(0) = \frac{1}{1-\gamma^\theta}\left(1 - g(0)^{-\theta}\right)$.
- $P(a < Z \le b) = \frac{1}{1-\gamma^\theta}\left(g(a+1)^{-\theta} - g(b+1)^{-\theta}\right)$.

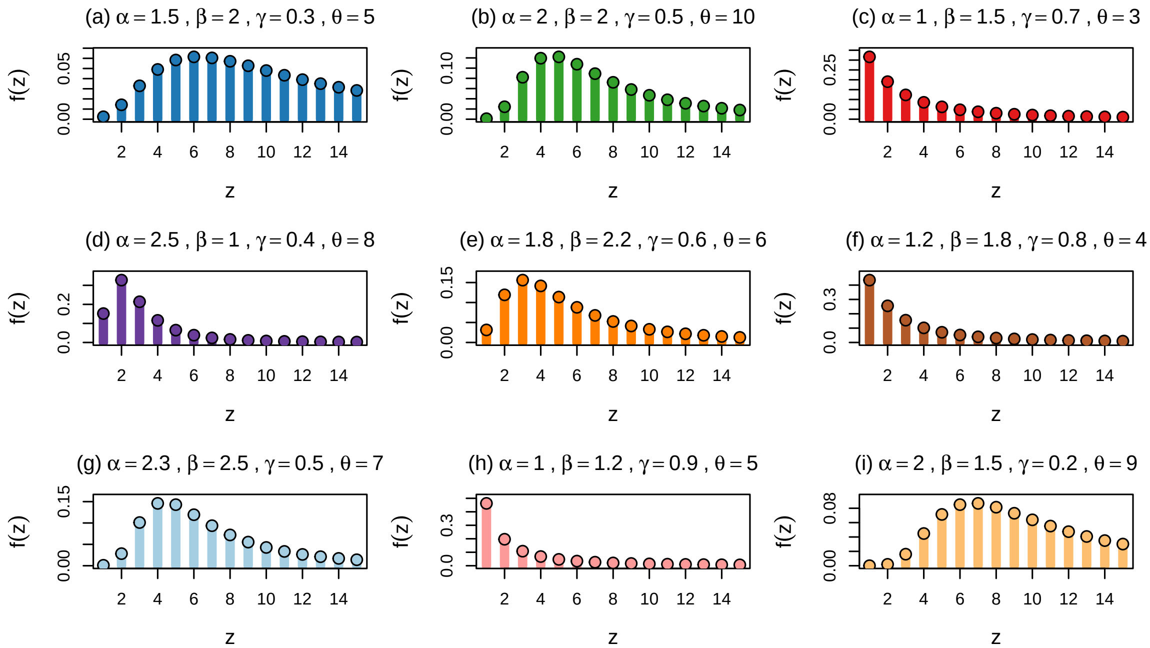

FIGURE 1. The PMF plots of DNGP distribution for different values of α, β, γ and θ

Result 3. *The survival function $S_Z(z)$ of DNGP distribution is given by*

$$S_Z(z) = \frac{1}{1-\gamma^\theta}\left(1 - g(z+1)^{-\theta}\right), \tag{6}$$

where $g(z) = \frac{z^\alpha + (\frac{1-\gamma}{\gamma})\beta^\alpha}{z^\alpha}$.

3. Statistical properties

In this section, we study the detailed properties of DNGP distribution.

Theorem 1. *DNGP distribution is heavy tailed.*

Proof. From survival function of the DNGP distribution is given in equation (6). We can simplify the term inside the bracket as follows:

$$\frac{(z+1)^{\alpha\theta}}{\left(\gamma(z+1)^\alpha + (1-\gamma)\beta^\alpha\right)^\theta} = \frac{(z+1)^{\alpha\theta}}{\left[\gamma(z+1)^\alpha\right]^\theta \left(1 + \frac{(1-\gamma)\beta^\alpha}{\gamma(z+1)^\alpha}\right)^\theta}$$

$$= \frac{1}{\gamma^\theta}\left(1 + \frac{c}{(z+1)^\alpha}\right)^{-\theta},$$

where $c = \dfrac{(1-\gamma)\beta^\alpha}{\gamma}$.

Substituting this expression into the survival function gives

$$S_Z(z) = \frac{1}{1-\gamma^\theta} - \frac{1}{1-\gamma^\theta}\left(1+\frac{c}{(z+1)^\alpha}\right)^{-\theta} = \frac{1-\left(1+\frac{c}{(z+1)^\alpha}\right)^{-\theta}}{1-\gamma^\theta}.$$

For large z, we have $\frac{c}{(z+1)^\alpha} \ll 1$. Using the first-order Taylor expansion $(1+u)^{-\theta} = 1-\theta u + O(u^2)$ for small u, we obtain

$$S_Z(z) \approx \frac{1-\left[1-\theta\frac{c}{(z+1)^\alpha}+O\left(\frac{1}{z^{2\alpha}}\right)\right]}{1-\gamma^\theta}$$

$$\sim \frac{\theta c}{1-\gamma^\theta}(z+1)^{-\alpha}.$$

Hence, for $z \to \infty$, $S_Z(z) \propto z^{-\alpha}$, which shows that the tail of the DNGP distribution decays according to a power law with index α.

Therefore, the DNGP distribution is heavy-tailed, since its survival function is regularly varying with index $-\alpha$. □

Remark 2. *In particular,*

$$S_Z(z) \sim K\, z^{-\alpha}, \quad K = \frac{\theta(1-\gamma)\beta^\alpha}{\gamma(1-\gamma^\theta)}.$$

Consequently, the moments exist only up to order α, that is,

$$E[Z^r] < \infty \quad \text{iff } r < \alpha.$$

Failure rates.

Result 4. *The failure rate of DNGP distribution is given by*

$$r_Z(z) = \frac{g(z+1)^{-\theta} - g(z)^{-\theta}}{1-g(z+1)^{-\theta}}. \tag{7}$$

Result 5. *The reverse failure rate of DNGP distribution is given by*

$$r_Z^*(z) = \frac{g(z+1)^{-\theta} - g(z)^{-\theta}}{g(z+1)^{-\theta} - \gamma^\theta}. \tag{8}$$

Moments.

Result 6. *The r^{th} moment of DNGP distribution is given by*

$$E(Z^r) = \sum_{z=[\beta]}^{\infty} z^r \frac{1}{1-\gamma^\theta}\left(g(z+1)^{-\theta} - g(z)^{-\theta}\right). \tag{9}$$

Theorem 2. *$E[Z^r] < \infty$ iff $r < \alpha$.*

Proof. For a nonnegative integer-valued random variable Z, one convenient representation for the r-th moment is

$$E[Z^r] = \sum_{z=0}^{\infty} z^r P(Z=z) = \sum_{k=1}^{\infty} \left[(k^r - (k-1)^r)\right] P(Z \geq k) = \sum_{k=1}^{\infty} \Delta_k(k^r)\, S_Z(k-1),$$

where $\Delta_k(k^r) = k^r - (k-1)^r$ and $S_Z(z)$ is given in equation (6). For large k a first-order expansion gives

$$\Delta_k(k^r) = k^r - (k-1)^r = r k^{r-1} + O(k^{r-2}).$$

Hence, using the asymptotic form of $S_Z(k-1) \sim K(k)^{-\alpha}$, the k-th term in the series behaves like

$$\Delta_k(k^r)\, S_Z(k-1) \sim (r k^{r-1}) \cdot (K k^{-\alpha}) = (rK)\, k^{r-1-\alpha}, \qquad (k \to \infty).$$

Thus the series for $E[Z^r]$ has the same convergence behaviour as the p-series $\sum_k k^{r-1-\alpha}$. That series converges iff its exponent is strictly less than -1, i.e.

$$r - 1 - \alpha < -1 \quad \Longleftrightarrow \quad r < \alpha.$$

Therefore $E[Z^r] < \infty$ iff $r < \alpha$. This completes the proof. □

Remark 3. *In particular*

- *The mean $E[Z]$ is finite iff $\alpha > 1$ ($r = 1$).*
- *The variance $\mathrm{Var}(Z)$ is finite iff the second moment is finite, i.e. iff $\alpha > 2$ ($r = 2$).*

4. Characterizations

This section deals with various characterizations of DNGP distribution. These characterizations are based on:(i) failure rate function and (ii) reverse failure rate function. Due to the nature of the DNGP distribution, we believe that our characterizations may be the only possible ones for this particular distribution.

Based on failure rate function.

Proposition 1. *If $Z \sim DNGP(\alpha, \beta, \gamma, \theta)$ iff its failure rate function satisfies the difference equation*

$$r_Z(k+1) - r_Z(k) = \frac{g(k+2)^{-\theta} - g(k+1)^{-\theta}}{1 - g(k+2)^{-\theta}} - \frac{g(k+1)^{-\theta} - g(k)^{-\theta}}{1 - g(k+1)^{-\theta}}; \quad k \in \mathbf{N_0} \tag{10}$$

with boundary condition

$$r_Z(0) = \frac{1}{g(1)^{\theta} - 1}.$$

Proof. If Z has PMF in equation (4) then clearly equation (10) holds. Now if equation (10) holds, then for every $Z \in \mathbf{N_0}$, we have

$$\sum_{k=0}^{z-1} r_Z(k+1) - r_Z(k) = \sum_{k=0}^{z-1} \frac{g(k+2)^{-\theta} - g(k+1)^{-\theta}}{1 - g(k+2)^{-\theta}} - \frac{g(k+1)^{-\theta} - g(k)^{-\theta}}{1 - g(k+1)^{-\theta}}.$$

$$r_Z(z) - r_Z(0) = \frac{g(z+1)^{-\theta} - g(z)^{-\theta}}{1 - g(z+1)^{-\theta}} - \frac{1}{g(1)^{\theta} - 1}.$$

In view of the fact $r_Z(0) = \frac{1}{g(1)^{\theta}-1}$, from the last equation we have

$$r_Z(z) = \frac{g(z+1)^{-\theta} - g(z)^{-\theta}}{1 - g(z+1)^{-\theta}},$$

which in view of equation (7), implies Z has PMF in equation (4). □

Based on reverse failure rate function.

Proposition 2. *If $Z \sim DNGP(\alpha, \beta, \gamma, \theta)$ iff its reverse failure rate function satisfies the difference equation*

$$r^*_Z(k+1) - r^*_Z(k) = \frac{g(k+2)^{-\theta} - g(k+1)^{-\theta}}{g(k+2)^{-\theta} - \gamma^{\theta}} - \frac{g(k+1)^{-\theta} - g(k)^{-\theta}}{g(k+1)^{-\theta} - \gamma^{\theta}} \quad k \in \mathbf{N_0} \tag{11}$$

with boundary condition $r^*_Z(0) = \frac{1}{1-\gamma^{\theta} g(1)^{\theta}}$.

Proof. If Z has PMF in equation (4) then clearly equation (11) holds. Now if equation (11) holds, then for every $Z \in \mathbf{N_0}$, we have

$$\sum_{k=0}^{z-1} r^*_Z(k+1) - r^*_Z(k) = \sum_{k=0}^{z-1} \frac{g(k+2)^{-\theta} - g(k+1)^{-\theta}}{g(k+2)^{-\theta} - \gamma^{\theta}} - \frac{g(k+1)^{-\theta} - g(k)^{-\theta}}{g(k+1)^{-\theta} - \gamma^{\theta}}.$$

Or,

$$r^*_Z(z) - r^*_Z(0) = \frac{g(z+1)^{-\theta} - g(z)^{-\theta}}{g(z+1)^{-\theta} - \gamma^{\theta}} - \frac{1}{1 - \gamma^{\theta} g(1)^{\theta}}.$$

In view of the fact that $r^*_Z(0) = \frac{1}{1-\gamma^{\theta} g(1)^{\theta}}$, from the last equation we have

$$r^*_Z(z) = \frac{g(z+1)^{-\theta} - g(z)^{-\theta}}{g(z+1)^{-\theta} - \gamma^{\theta}},$$

which in view of equation (8), implies Z has PMF in equation (4). □

5. DNGP regression model

In this section, we formulate a regression model for a count response variable assumed to follow the DNGP distribution.

Suppose Z is a count response variable that follows DNGP distribution given in equation (4) and is associated with a set of covariates. Here we want to fit the response variable by using some covariates. Let $X_i = (x_{i1}, x_{i2}, x_{i3}, ...x_{ik})$ be a k dimensional vector of explanatory varaibles. The mean of DNGP distribution is not in a closed form. Since the distribution is characterized by a parameter γ that lies between 0 and 1, we assume that the parameter is a function of x_i given by $\gamma(x_i) = \phi(x_i, \lambda)$, where $0 < \phi(x_i, \lambda) < 1$ is a known function of x_i and k dimensional column vector $\lambda = (\lambda_0, \lambda_1, \lambda_2, ..., \lambda_k)'$ of regression parameters. The logit function is

$$\phi(x_i, \lambda) = \gamma(x_i) = \frac{e^{x_i'\lambda}}{1 + e^{x_i'\lambda}},$$

since $0 < \gamma < 1$.

The DNGP regression model is defined by using PMF given in equation (4) with the function

$$\gamma(x_i) = \frac{e^{x_i'\lambda}}{1 + e^{x_i'\lambda}}. \tag{12}$$

This results in the DNGP regression model expressed as

$$f(z_i) = P(Z = z_i \mid x_i) = \frac{1}{1 - (\gamma(x_i))^{\theta}} \left(g_i(z_i + 1)^{-\theta} - g_i(z_i)^{-\theta} \right]), \tag{13}$$

where, $g_i(y) = \frac{z_i^{\alpha} + \left(\frac{1-\gamma(x_i)}{\gamma(x_i)}\right)\beta^{\alpha}}{z_i^{\alpha}}$, $\alpha > 0$, $\beta > 0$, $0 < \gamma(x_i) < 1$, $\theta > 0$ and $\gamma(x_i)$ given in equation (12).

In the DNGP regression model in equation (13), one can also assume that the parameters α, β and θ are functions of covariates. Since the parameters $\alpha > 0$, $\beta > 0$ and $\theta > 0$, we can assume that the function $\alpha(x_i) = \beta(x_i) = \theta(x_i) = exp(x_i'\lambda)$. In this article, we assume that the parameters α, β and θ are nuisance parameters. There are many regression models, like Poisson regression model, the mean is taken as the function of covariates. But in the case of DNGP mean is not in closed form and it is a function (α, β, γ, θ). So we take the parameter γ as the function of covariates. Therefore, the mean of DNGP regression model is an implicit function of covariates since the mean of DNGP distribution is a function of γ.

Fitted median. The relationship between certain explanatory variables and the median can also be obtained. The median is more appropriate for count data. Also, the median has a closed form for the DNGP model.

$$Median = M = \beta[(\frac{1-\gamma}{\gamma})\frac{1}{[1 - \frac{1-\gamma^{\theta}}{2}]^{-1/\theta} - 1}]^{1/\alpha} - 1. \tag{14}$$

$$M + 1 = \beta[(\frac{1-\gamma}{\gamma})\frac{1}{[1 - \frac{1-\gamma^{\theta}}{2}]^{-1/\theta} - 1}]^{1/\alpha}. \tag{15}$$

Substituting $\gamma = \frac{e^{x_i'\lambda}}{1+e^{x_i'\lambda}}$ in equation (15), we have,

$$M + 1 = \beta[e^{-x_i'\lambda}\frac{1}{(1/2 - (e^{-x_i'\lambda} + 1)^{-\theta})^{-1/\theta}} - 1]^{1/\alpha}. \tag{16}$$

Taking logarithm on equation (16), we get,

$$\log(M+1) = \log(\beta) - \frac{x_i'\lambda}{\alpha} - \log((1/2 - (e^{-x_i'\lambda} + 1)^{-\theta})^{-1/\theta} - 1)). \tag{17}$$

6. Estimation

The estimation of the unknown parameters is done using the MLE method. Let $(x_i, z_i), i = 1, 2, ..., n$ be a sample of 'n' independent observations from the PMF. The likelihood function based on this observed sample is given by,

$$L = f(z_i|x_i) = \prod_{i=1}^{n} \frac{1}{1 - \gamma(x_i)^{\theta}} \times \left(g_i(z_i + 1)^{-\theta} - g_i(z_i)^{-\theta} \right]), \tag{18}$$

where $\gamma_{x_i} = \gamma_i = \frac{e^{x_i'\lambda}}{1+e^{x_i'\lambda}}$.

The log-likelihood function of the DNGP regression model is obtained by taking the natural logarithm of the likelihood function:

$$\begin{aligned} l &= \sum_{i=1}^{n} \log \left[\frac{1}{1-[\gamma(x_i)]^{\theta}} \left(g_i(z_i+1)^{-\theta} - g_i(z_i)^{-\theta} \right) \right] \\ &= -\sum_{i=1}^{n} \log \left(1-[\gamma(x_i)]^{\theta} \right) + \sum_{i=1}^{n} \log \left(g_i(z_i+1)^{-\theta} - g_i(z_i)^{-\theta} \right), \end{aligned} \tag{19}$$

where

$$\gamma(x_i) = \frac{e^{x_i'\lambda}}{1+e^{x_i'\lambda}}, \quad g_i(z) = \frac{z^{\alpha} + \left(\frac{1-\gamma(x_i)}{\gamma(x_i)}\right)\beta^{\alpha}}{z^{\alpha}}.$$

Then to obtain the ML estimates, the first partial derivatives w.r.t each unknown parameter are obtained and set equal to zero.

Define

$$\gamma_i = \gamma(x_i) = \frac{e^{x_i^{\top}\lambda}}{1+e^{x_i^{\top}\lambda}}, \qquad g_i(y) = 1 + \frac{1-\gamma_i}{\gamma_i}\frac{\beta^{\alpha}}{y^{\alpha}},$$

and $u_i = g_i(z_i+1)^{-\theta} - g_i(z_i)^{-\theta}$.

The log–likelihood becomes

$$\ell = -\sum_{i=1}^{n} \log(1-\gamma_i^{\theta}) + \sum_{i=1}^{n} \log u_i.$$

Using $\dfrac{\partial g_i(z)}{\partial \alpha} = (g_i(z)-1)(\log\beta - \log z)$, we obtain

$$\begin{aligned} \frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \frac{1}{u_i} \Big\{ & -\theta\, g_i(z_i+1)^{-(\theta+1)} \big(g_i(z_i+1)-1\big) \big(\log\beta - \log(z_i+1)\big) \\ & + \theta\, g_i(z_i)^{-(\theta+1)} \big(g_i(z_i)-1\big) \big(\log\beta - \log z_i\big) \Big\}. \end{aligned} \tag{20}$$

Using $\dfrac{\partial g_i(z)}{\partial \beta} = (g_i(z)-1)\dfrac{\alpha}{\beta}$, we get

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} \frac{1}{u_i} \left\{ -\theta \frac{\alpha}{\beta}\, g_i(z_i+1)^{-(\theta+1)} \big(g_i(z_i+1)-1\big) + \theta \frac{\alpha}{\beta}\, g_i(z_i)^{-(\theta+1)} \big(g_i(z_i)-1\big) \right\}. \tag{21}$$

Using $\dfrac{\partial \gamma_i}{\partial \lambda} = \gamma_i(1-\gamma_i)x_i$ and

$$\frac{\partial g_i(y)^{-\theta}}{\partial \gamma_i} = \theta\, g_i(y)^{-(\theta+1)} \big(g_i(y)-1\big) \frac{1}{(1-\gamma_i)\gamma_i},$$

$$\begin{aligned} \frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} \Bigg[& \frac{\theta\, \gamma_i^{\theta-1}}{1-\gamma_i^{\theta}}\, \gamma_i(1-\gamma_i) \\ & + \frac{\theta}{u_i} \Big(g_i(z_i+1)^{-(\theta+1)} \big(g_i(z_i+1)-1\big) - g_i(z_i)^{-(\theta+1)} \big(g_i(z_i)-1\big) \Big) \Bigg] x_i. \end{aligned} \tag{22}$$

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{n} \frac{\gamma_i^{\theta} \log \gamma_i}{1-\gamma_i^{\theta}} + \sum_{i=1}^{n} \frac{-\log\big(g_i(z_i+1)\big)\, g_i(z_i+1)^{-\theta} + \log\big(g_i(z_i)\big)\, g_i(z_i)^{-\theta}}{u_i}. \tag{23}$$

Setting the four expressions given in equations (20)-(23) equal to zero yields the ML equations to be solved numerically. These system of likelihood equations cannot be solved simultaneously and hence they do not have an analytic solution. Therefore MLEs of α, β, λ and θ can be found by directly maximizing the loglikelihood function numerically. This can be performed easily using an iterative numerical optimization tool, such as the `optim` function in **R**.

Given that the parameter inferences are performed using the ML method, then under some regularity conditions, these estimators have standard asymptotic properties. That is, the MLEs $\hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{\theta}$ have some characteristics as follows:

i) They are asymptotically consistent(unbiased).

ii)They have asymptotically variance-covariance matrix obtained from the inverse of the expected Fisher information matrix,

$$I = \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\theta} & I_{\alpha\lambda} \\ I_{\beta\alpha} & I_{\beta\beta} & I_{\beta\theta} & I_{\beta\lambda} \\ I_{\theta\alpha} & I_{\theta\beta} & I_{\theta\theta} & I_{\theta\lambda} \\ I_{\lambda\alpha} & I_{\lambda\beta} & I_{\lambda\theta} & I_{\lambda\lambda} \end{bmatrix}$$

where $I_{\eta_i\eta_j} = -E[\frac{\partial^2 l}{\partial_{\eta_i\eta_j}}], i, j = 1, 2, 3, 4.$

iii) The MLEs $\hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{\theta}$ ae asymptotically normally distributed.

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\beta}-\beta \\ \hat{\theta}-\theta \\ \hat{\lambda}-\lambda \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, I^{-1}, \right)$$

where α, β, θ and λ are true value of the parameters and I is the Fisher information matrix.

Simulation study. The DNGP regression model is generated using $\gamma(x_i) = exp(\lambda_0+\lambda_1 x_i)/[1+exp(\lambda_0+\lambda_1 x_i)]$, where x_i is a random sample from uniform(0,1) distribution. We take different sample sizes n=50, 100, 150 and different values for parameter α, β, θ, λ_0 and λ_1. Each simulation is based on 1000 replications. The covariate x_i is fixed throughout the simulation study. We consider the values $\alpha = 2.0$, $\beta = 1.0$, $\lambda_0 = 1.2$, $\lambda_1 = 0.5$ and $\theta = 0.1$ for under-dispersed data, while the values $\alpha = 1.5$, $\beta = 1.0$, $\lambda_0 = 1.2$, $\lambda_1 = 0.5$ and $\theta = 0.1$ for over-dispersed data. The ML estimates and MSE of the parameters were obtained.

TABLE 1. Regression parameter estimates of α, β, λ_0, λ_1 and θ

sample size (n)	α	β	λ_0	λ_1	θ	$\hat{\alpha}$ $MSE(\hat{\theta})$	$\hat{\beta}$ $MSE(\hat{\beta})$	$\hat{\lambda}_0$ $MSE(\hat{\lambda}_0)$	$\hat{\lambda}_1$ $MSE(\hat{\lambda}_1)$	$\hat{\theta}$ $MSE(\hat{\theta})$
50						2.152	1.359	1.280	0.539	0.166
						(0.227)	(0.289)	(0.168)	(0.170)	(0.153)
100	2.0	1.0	1.2	0.5	0.1	2.047	1.132	1.271	0.522	0.146
						(0.139)	(0.215)	(0.157)	(0.127)	(0.193)
150						2.037	1.054	1.003	0.520	0.122
						(0.126)	(0.110)	(0.120)	(0.119)	(0.244)
50						1.512	0.877	1.433	0.553	0.083
						(0.156)	(0.151)	(0.544)	(0.833)	(0.740)
100	1.5	1.0	1.2	0.5	0.1	1.510	1.062	1.331	0.552	0.085
						(0.151)	(0.109)	(0.412)	(0.501)	(0.687)
						[0.720]	[0.621]	[0.507]	[0.457]	[0.852]
150						1.501	1.057	1.212	0.551	0.122
						(0.112)	(0.101)	(0.109)	(0.200)	(0.127)

Table 1 lists the simulation results for paramater vectors given above. The simulation results are evaluated by means of averages of the estimates and MSEs. From the Table 1, we see that estimates are approaches to the parameter values and MSEs decrease as sample size n increases.

7. Numerical illustration

7.1. **Heart failure clinical records data:** This dataset is taken from the medical records of 299 heart failure patients collected at the Faisalabad Institute of Cardiology and at the Allied Hospital in Faisalabad (Punjab, Pakistan), during April–December 2015. For more details see ([2]). The response variable is chosen as the ejection fraction which is the the percentage of how much blood the left ventricle pumps out with each contraction. The covarites are taken as age of the patients, level of serum creatinine in the blood and level of serum sodium in the blood. The mean of the response variable is 38.0836 and variance is 140.0635.

We have fitted the DNGP regression model to the data and compare with other popular count data regression models such as Poisson and negative binomial distributions. The estimated value of the parameters along with AIC values are given in Table 2.

Table 2. Fitting with other regression models

	Intercept	level of serum creatinine	other	AIC
DNGPR	39.753	-55.241	$\hat{\alpha}$ =18.697 $\hat{\beta}$=74.952 $\hat{\theta} = 1.041 \times 10^{-5}$	1055.523
Poisson	3.645	-0.003		2709.400
Negative Binomial	3.644	-0.003	k=14.495	2312.600

To compare the relative fit of two competing, non-nested models, [16] proposed Vuong test. This test allows for direct comparison of models that are not nested within each other, such as a Poisson regression versus DP regression model. The Vuong statistic is given by:

$$V = \frac{\sqrt{n}\bar{d}}{s_d},$$

where, $\bar{d}$ is the mean log-likelihood difference between two non nested models, s_d is the standard deviation of differences and n is the sample size.
Under the null hypothesis that both models are observationally equivalent (i.e., have equal expected K-L divergence from the true model), the statistic V is asymptotically standard normal: $V \sim N(0, 1)$. The hypothesis considered are
H_0: Both models fit the data equally well.
H_1:One model fits significantly better than the other.

The decision taken at 5% significance level if:

- $V > 1.96$: Model 1 provides a significantly better fit.
- $V < -1.96$: Model 2 provides a significantly better fit.
- $|V| \leq 1.96$: No significant difference between the models.

Table 3. Vuong test results for heart failure data

	statistic	p value
DNGPR & Poisson	4.7361	0.00702
DNGPR& Negative binomial	4.1328	0.000697

From Table 3, it is clear that p value is less than $\alpha* = 0.05$, we reject our null hypothesis. The model DNGPR is preferred over other two models.

8. Conclusions

This paper introduced and examined a regression model based on the discrete new generalized Pareto (DNGP) distribution for analyzing count data. The proposed model is flexible enough to handle both underdispersed and overdispersed data, making it a useful alternative to classical count regression models. Although closed-form expressions for the mean and variance are not available, the availability of closed-form probability mass and cumulative distribution functions simplifies probability computations. Real data applications demonstrate that the DNGP regression model provides a better fit than Poisson and negative binomial regression models. A limitation of the proposed approach is the computational complexity of the likelihood function, which may lead to convergence issues and sensitivity to initial values during parameter estimation.

Competing interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] A. Buddana, T.J. Kozubowski, Discrete Pareto Distributions, Econ. Qual. Control. 29 (2014), 143–156. https://doi.org/10.1515/eqc-2014-0014.

[2] D. Chicco, G. Jurman, Machine Learning Can Predict Survival of Patients with Heart Failure from Serum Creatinine and Ejection Fraction Alone, BMC Med. Inform. Decis. Mak. 20 (2020), 16. https://doi.org/10.1186/s12911-020-1023-5.

[3] E.L. Frome, M.H. Kutner, J.J. Beauchamp, Regression Analysis of Poisson-Distributed Data, J. Am. Stat. Assoc. 68 (1973), 935–940. https://doi.org/10.1080/01621459.1973.10481449.

[4] I. Ghosh, A New Discrete Pareto Type (IV) Model: Theory, Properties and Applications, J. Stat. Distrib. Appl. 7 (2020), 3. https://doi.org/10.1186/s40488-020-00104-x.

[5] K. Jayakumar, J. Jose, Discrete New Generalized Pareto Distribution, Statistica 82 (2022), 373–391. https://doi.org/10.6092/issn.1973-2201/11380.

[6] K. Jayakumar, J. Jose, On an Alternative to Discrete Pareto Model, Aust. J. Stat. 54 (2025), 51–78.

[7] K. Jayakumar, B. Krishnan, G.G. Hamedani, On a New Generalization of Pareto Distribution and Its Applications, Commun. Stat. - Simul. Comput. 49 (2018), 1264–1284. https://doi.org/10.1080/03610918.2018.1494281.

[8] D.W. Jorgenson, Multiple Regression Analysis of a Poisson Process, J. Am. Stat. Assoc. 56 (1961), 235–245. https://doi.org/10.1080/01621459.1961.10482106.

[9] J. Jose, K. Jayakumar, On Some Aspects of Generalized Discrete Pareto Distribution, Asian J. Stat. Sci. 5 (2025), 31–56.

[10] T.J. Kozubowski, A.K. Panorska, M.L. Forister, A Discrete Truncated Pareto Distribution, Stat. Methodol. 26 (2015), 135–150. https://doi.org/10.1016/j.stamet.2015.04.002.

[11] H. Krishna, P. Singh Pundir, Discrete Burr and Discrete Pareto Distributions, Stat. Methodol. 6 (2009), 177–188. https://doi.org/10.1016/j.stamet.2008.07.001.

[12] P. McCullagh, J. Nelder, Generalized Linear Models, Chapman and Hall, (1989). https://doi.org/10.1201/9780203753736.

[13] J.A. Nelder, R.W.M. Wedderburn, Generalized Linear Models, J. R. Stat. Soc. Ser. A Gen. 135 (1972), 370–384. https://doi.org/10.2307/2344614.

[14] P. BA, On Discrete Three Parameter Burr Type XII and Discrete Lomax Distributions and Their Applications to Model Count Data from Medical Science, Biom. Biostat. Int. J. 4 (2016), 1–15. https://doi.org/10.15406/bbij.2016.04.00092.

[15] F. Prieto, E. Gómez-Déniz, J.M. Sarabia, Modelling Road Accident Blackspots Data with the Discrete Generalized Pareto Distribution, Accid. Anal. Prev. 71 (2014), 38–49. https://doi.org/10.1016/j.aap.2014.05.005.

[16] Q.H. Vuong, Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses, Econometrica 57 (1989), 307–333. https://doi.org/10.2307/1912557.