Eur. J. Stat. 5 (2025) 15
doi: 10.28924/ada/stat.5.15

Discrete Harris Extended Inverse Exponential Distribution and Its Applications to COVID-19 Data

Jabir Bengalath*, Bindu Punathumparambath

Department of Statistics, Govt. Arts and Science College, Calicut, University of Calicut, Kerala, India
jabirbengu@gmail.com, ppbindukannan@gmail.com
**Correspondence: jabirbengu@gmail.com*

Abstract. In this paper, we introduce a novel discrete probability distribution, referred to as the discrete Harris extended inverse exponential distribution, and explore its fundamental properties. We demonstrate that the proposed model serves as a generalization of the discrete Marshall-Olkin inverted exponential distribution. Key theoretical characteristics of the proposed model are derived, including its probability generating function, moments, hazard rate function, cumulative hazard rate function, reversed hazard rate function, mean residual life, and quantile function. Parameter estimation is performed using various estimation techniques, including the maximum likelihood estimation, Anderson-Darling, Cramér-von Mises, ordinary least squares, and weighted least squares. In addition, a simulation study is conducted to evaluate the performance of the estimators. The applicability of the proposed distribution is further illustrated through its fit to two discrete real-world data sets of COVID-19 from China and Pakistan.

1. Introduction

In scientific studies, discrete variables are common, particularly in reliability or life testing experiments where measuring a device's lifespan on a continuous scale is often impractical. For example, the durability of an on/off switch is typically quantified by counting its operational cycles, making its lifespan a discrete random variable. In survival analysis, researchers may record the number of days a patient survives post-treatment or the time from remission to relapse, relying on counts rather than continuous measurements. By discretizing continuous distributions, researchers have developed various discrete lifetime distributions. These include the discrete Weibull distribution introduced by Nakagawa and Osaki [1], a second type by Stein and Dattero [2], a third type by Padgett and Spurrier [3], the discrete exponential distribution by Sato et al. [4], the discrete

Received: 24 Jun 2025.
Key words and phrases. Harris extended inverse exponential distribution; Discrete Harris extended inverse exponential distribution; Statistical properties; Estimation; Simulation; Application.

normal distribution by Roy [5], the discrete Rayleigh distribution by Roy [6], the discrete Laplace distribution by Inusah and Kozubowski [7], the discrete skew-Laplace distribution by Kozubowski and Inusah [8], the discrete Burr and Pareto distributions by Krishna and Pundir [9], and the discrete inverse Weibull distribution by Jazi et al. [10], among others, are notable contributions to the literature. Discretization is vital in variable selection techniques, extending beyond the mere transformation of continuous variables into discrete forms. It can significantly improve the efficacy of classification algorithms applied to high-dimensional biomedical data analysis. When constructing a discrete counterpart of a continuous distribution, preserving key properties of the original distribution is often a priority. The literature presents various approaches to discretization, as explored by different authors, including Bracquemond and Gaudoin [11] and Chakraborty [12].

In this study, we introduce a novel discrete Harris extended inverted distribution, termed the discrete Harris extended inverse exponential distribution. This family enhances modeling flexibility for non-monotonic and heavy-tailed datasets. Aly and Benkherouf [13] generalized the Marshall-Olkin (MO) distribution to introduce a more flexible model using the pgf of the Harris distribution (Harris, 1948) for generating new distributions. Cordeiro et al. [14] introduced Harris extended Lindley distribution. Tomy et al. [15] Harris extended modified Lindley. Jabir and Bindu [16, 20] proposed the Harris extended inverted Kumaraswamy and Harris extended Dagum distributions. Sophia et al. [17] introduced the discrete Harris extended Weibull distribution. The survival function of Harris family of distribution is given by

$$\bar{F}_{HE}(x) = \frac{\alpha^{\frac{1}{\beta}}\bar{G}(x)}{\left[1-\bar{\alpha}\bar{G}(x)^{\beta}\right]^{\frac{1}{\beta}}}, \quad x > 0, \tag{1}$$

where $\bar{\alpha} = 1-\alpha$, $\alpha > 0$, and $\beta > 0$. The newly introduced parameters α and β represent additional shape parameters that allow for enhanced flexibility. The survival function and probability density function of the inverse exponential distribution are provided below.

$$\bar{G}_{IE}(x) = 1 - e^{-\frac{\theta}{x}}, \quad x > 0,\ \theta > 0, \tag{2}$$

and

$$g_{IE}(x) = \frac{1}{\theta x^2}e^{-\frac{\theta}{x}}, \quad x > 0,\ \theta > 0. \tag{3}$$

The remaining part of the paper is organized in the following order : Section 2 describes the Harris extended inverse exponential distribution. In Section 3, we defined the discrete Harris extended inverse exponential distribution with its associated statistical properties, such as quantile function, median, skewness and kurtosis, moments and moment generating function. The estimation of the parameters by using different estimation methods such as maximum likelihood, Anderson-Darling, Ordinary Least Squares, Weighted Least Squares and Cramer-von Mises, is discussed in Section 4.

In Section 5, certain simulated data sets are provided to understand the behavior of the proposed model in a large sample. Two real datasets related to COVID-19 are analyzed using the proposed distribution in Section 6. Finally, the study is concluded in Section 7.

2. Harris extended inverse exponential distribution

Definition 1. *A random variable X is said to have Harris extended inverse exponential distribution (HEIE) with parameters θ, α, β, denoted by $X \sim HEIE(\theta, \alpha, \beta)$ if its probability density function (pdf) is given by,*

$$g_{HEIE}(x) = \frac{\alpha^{\frac{1}{\beta}} \frac{1}{\theta x^2} e^{-\frac{\theta}{x}}}{\left(1 - \bar{\alpha}\left(1 - e^{-\frac{\theta}{x}}\right)^{\beta}\right)^{\left(1+\frac{1}{\beta}\right)}}, \quad x > 0, \tag{4}$$

where $\theta > 0$, $\bar{\alpha} = 1 - \alpha$, $\alpha > 0$ and $\beta > 0$.

Definition 2. *If $X \sim HEIE(\theta, \alpha, \beta)$ then its cumulative distribution function (cdf) is given by,*

$$F_{HEIE}(x) = 1 - \frac{\alpha^{\frac{1}{\beta}}\left(1 - e^{-\frac{\theta}{x}}\right)}{\left(1 - \bar{\alpha}\left(1 - e^{-\frac{\theta}{x}}\right)^{\beta}\right)^{\frac{1}{\beta}}}, \quad x > 0, \tag{5}$$

where $\theta > 0$, $\bar{\alpha} = 1 - \alpha$, $\alpha > 0$ and $\beta > 0$.

- The general approach of discretizing a continuous variable is to introduce a floor function or greatest integer function of X i.e., $[X]$ (the greatest integer less than or equal to X till it reaches the integer), in order to introduce grouping on a time axis.
- A discrete HEIE variable X_d, can be viewed as the discrete concentration of the continuous HEIE variable X. where the corresponding probability mass function of X_d can be written as:
$$P(X_d = x) = p(x) = S(x) - S(x+1), \quad x > 0, \tag{6}$$
where, $S(x)$ is the survival function of HEIE distribution.

3. discrete Harris extended inverse exponential distribution

Definition 3. *If $X_d \sim DHEIE(\alpha, \beta, q)$ then its probability mass function (pmf) and cumulative distribution function (cdf) are given by,*

$$p(x) = \alpha^{\frac{1}{\beta}} \left(\frac{1 - q^{\frac{1}{x}}}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x}}\right)^{\beta}\right)^{\frac{1}{\beta}}} - \frac{1 - q^{\frac{1}{x+1}}}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}} \right), \quad x = 0, 1, 2, \ldots \tag{7}$$

and

$$F(x) = 1 - \frac{\alpha^{\frac{1}{\beta}} \left(1 - q^{\frac{1}{x+1}}\right)}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}, \quad x = 0, 1, 2, \ldots \tag{8}$$

where $q = e^{-\theta}$, $0 < q < 1$, $\bar{\alpha} = 1 - \alpha$, $\alpha > 0$ *and* $\beta > 0$.

- When $\beta = 1$ the distribution reduces to discrete Marshall-Olkin inverse exponential (DMOIE) distribution.

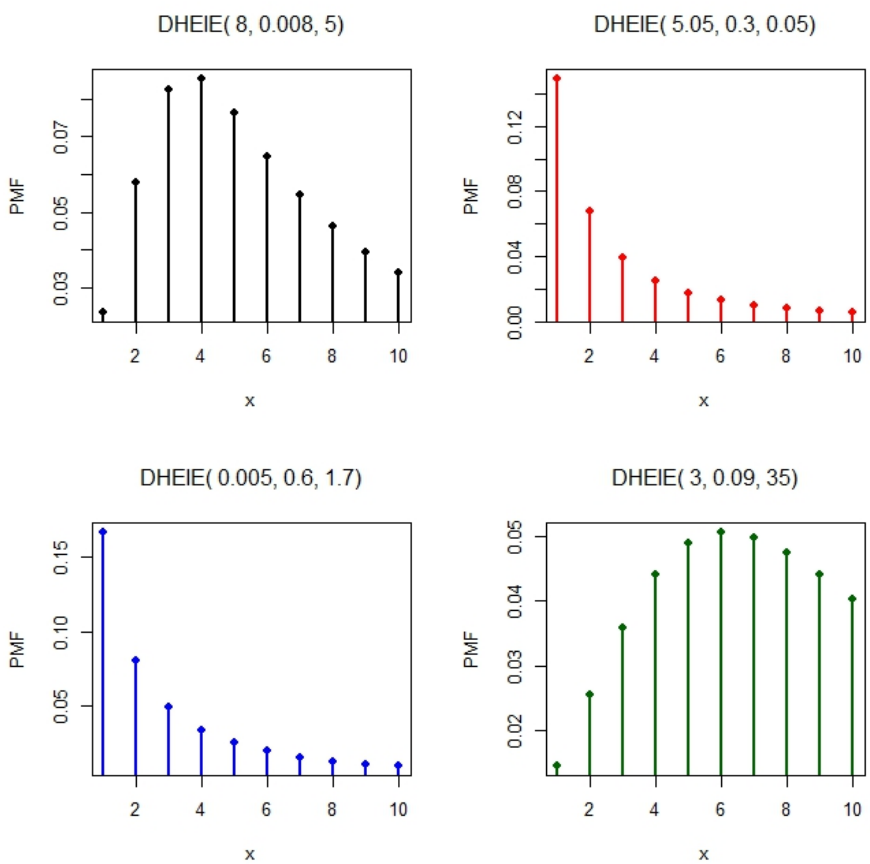

Figure 1. Plot of pmf of DHEIE for various values of α, β and q.

3.1. **Survival function.**

Definition 4. *The survival function (sf) of the DHEIE distribution is given by,*

$$S(x) = \frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}, \quad x = 0, 1, 2, \ldots \tag{9}$$

where $q = e^{-\theta}$, $0 < q < 1$, $\bar{\alpha} = 1-\alpha$, $\alpha > 0$ *and* $\beta > 0$.

3.2. **Hazard function.**

Definition 5. *The failure rate function (or hazard rate function) (hrf),* $h(x) = \frac{p(x)}{S(x)}$ *of the DHEIE distribution is given by,*

$$h(x) = \frac{\left(1-q^{1/x}\right)\left(1-\bar{\alpha}(1-q^{1/(x+1)})^{\beta}\right)^{1/\beta}}{\left(1-q^{1/(x+1)}\right)\left(1-\bar{\alpha}(1-q^{1/x})^{\beta}\right)^{1/\beta}} - 1, \quad x = 0, 1, 2, \ldots \tag{10}$$

where $q = e^{-\theta}$, $0 < q < 1$, $\bar{\alpha} = 1-\alpha$, $\alpha > 0$ *and* $\beta > 0$.

Figures 1 and 2 depict the pmf and hrf of the DHEIE distribution for varying parameter values. As illustrated, The pmf of the DHEIE distribution is a versatile statistical tool capable of modeling a wide range of discrete data behaviors. As illustrated across the plots, the distribution can take on different shapes depending on its parameters ranging from sharply right-skewed to symmetric or even slightly left-skewed forms. It can capture unimodal patterns where the probability mass is concentrated around a central value, as well as distributions where the mass is heavily skewed toward one end of the support. This flexibility allows the DHEIE distribution to accurately represent real-world scenarios where data may be clustered, dispersed, or asymmetrically distributed. The ability to control the peak location, spread, and skewness through its parameters makes it particularly well-suited for applications involving complex, non-uniform data, such as those found in environmental monitoring, resource usage, or other sustainability-related fields. Its adaptability enhances the modeler's ability to capture true variability in observed phenomena, supporting more precise analysis and more reliable decision-making.

Furthermore, the hrf of the DHEIE distribution is a powerful tool for modeling time-to-event data across diverse applications. This function quantifies the conditional probability of an event occurring at a specific time, given it has not yet occurred. As shown in the accompanying plots, the hrf of the DHEIE distribution can exhibit decreasing or upside-down bathtub (UBT) shapes, depending on parameter values. This versatility enables the DHEIE distribution to capture a wide range of real-world phenomena. A decreasing hrf is well-suited for scenarios where the risk of an event-such as failure-diminishes over time, as seen in reliability studies of robust systems like industrial machinery. In contrast, a UBT-shaped hrf, characterized by an initially high risk that

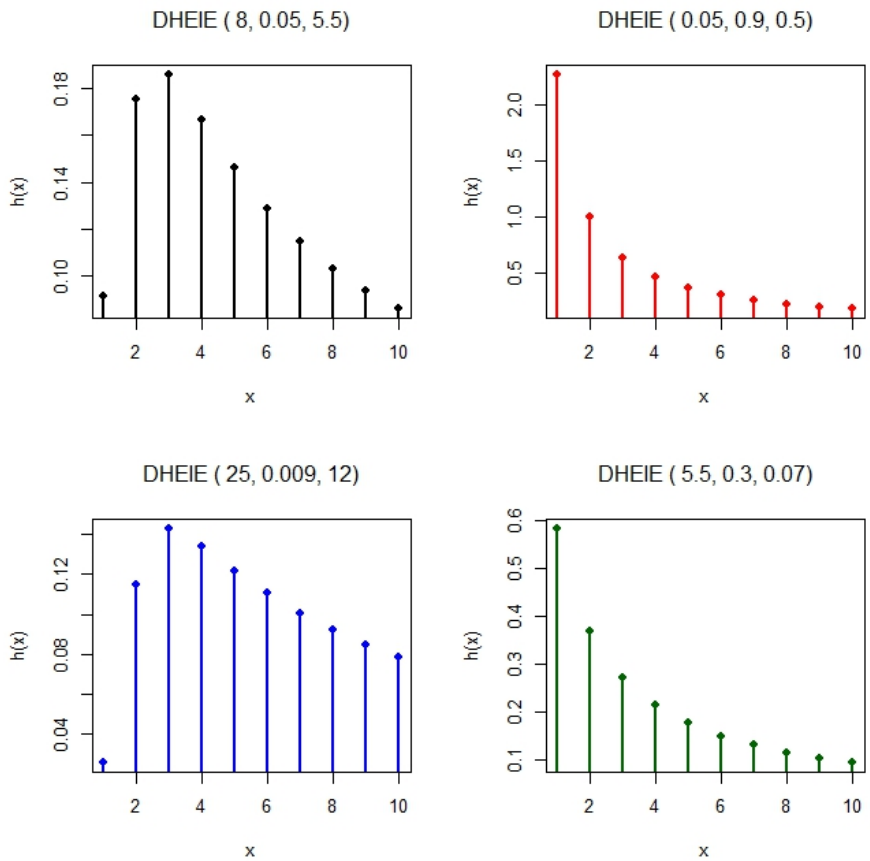

FIGURE 2. Plot of the hrf of DHEIE for various values of α, β, and q.

decreases to a minimum before rising again, is ideal for modeling systems with early vulnerabilities followed by a stable period and later wear-out, such as electronic components or certain biological processes. The flexibility of the DHEIE hrf makes it particularly valuable for sustainability-focused modeling, where understanding the timing and risk of events—such as equipment failure, resource depletion, or environmental incidents—is critical. By accurately capturing these dynamic risk profiles, the DHEIE hrf supports precise predictions and informed decision-making in domains like engineering, environmental management, and resource planning.

3.3. **Cumulative hazard function.**

Definition 6. *The cumulative hazard rate function (chf) of the DHEIE distribution is given by,*

$$H(x) = \sum_{t=0}^{x} h(t) = \sum_{t=0}^{x} \frac{\left(1-q^{1/t}\right)\left(1-\bar{\alpha}(1-q^{1/(t+1)})^{\beta}\right)^{1/\beta}}{\left(1-q^{1/(t+1)}\right)\left(1-\bar{\alpha}(1-q^{1/t})^{\beta}\right)^{1/\beta}} - 1, \quad x = 0, 1, 2, \ldots$$

where $q = e^{-\theta}$, $0 < q < 1$, $\bar{\alpha} = 1-\alpha$, $\alpha > 0$ *and* $\beta > 0$.

3.4. Mean residual life function.

Definition 7. *The mean residual life function, $L(x) = \mathbb{E}\left[(X-x)/X \geq x\right]$ of the DHEIE distribution is given by,*

$$L(x) = \sum_{j=x}^{\infty} \prod_{i=x}^{j} \left[2 - \frac{\left(1-q^{1/i}\right)\left(1-\bar{\alpha}(1-q^{1/(i+1)})^{\beta}\right)^{1/\beta}}{\left(1-q^{1/(i+1)}\right)\left(1-\bar{\alpha}(1-q^{1/i})^{\beta}\right)^{1/\beta}} \right], \quad x = 0, 1, 2, \ldots$$

where $q = e^{-\theta}$, $0 < q < 1$, $\bar{\alpha} = 1-\alpha$, $\alpha > 0$ and $\beta > 0$.

3.5. Reverse hazard rate function.

Definition 8. *The Reverse hazard rate function, $h^{*}(y) = P(Y = y \mid Y \leq y)$ of the DHEIE distribution is given by,*

$$h^{*}(x) = \frac{\alpha^{\frac{1}{\beta}} \left[\left(1-q^{\frac{1}{x+1}}\right) \left(1+\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{-\frac{1}{\beta}} - \left(1-q^{\frac{1}{x}}\right)\left(1+\bar{\alpha}\left(1-q^{\frac{1}{x}}\right)^{\beta}\right)^{-\frac{1}{\beta}} \right]}{1+\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)\left(1+\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{-\frac{1}{\beta}}},$$

where $q = e^{-\theta}$, $0 < q < 1$, $\bar{\alpha} = 1-\alpha$, $\alpha > 0$ and $\beta > 0$.

3.6. **Quantile function and median.** The DHEIE distribution can be simulated using inverse cdf method,

$$Q(u) = \lceil x_u \rceil = \left\lceil \left\{ \frac{\log\left[1-(1-u)\left(\alpha+\bar{\alpha}\,(1-u)^{\beta}\right)^{-1/\beta}\right]^{-1}}{\log q} \right\} - 1 \right\rceil, \tag{11}$$

where $\lceil x_u \rceil$ denotes the smallest integer greater than or equal to x_u and $0 < u < 1$. By utilizing equation (13), we can compute the first and third quartiles by substituting $q = 0.25$ and $q = 0.75$, respectively.

In particular, median is given by,

$$Q(u) = \lceil x_u \rceil = \left\lceil \left\{ \frac{\log\left[1-0.5\left(\alpha+\bar{\alpha}\,(0.5)^{\beta}\right)^{-1/\beta}\right]^{-1}}{\log q} \right\} - 1 \right\rceil.$$

3.7. **Skewness and kurtosis.** By utilizing equation (13), we can compute Galton's skewness (S_k) and Moor's Kurtosis (K_r) using the formulae given by

$$S_k = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)},$$

and the measure of Kurtosis, K_r

$$K_r = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)},$$

3.8. **Moments.** Let $X \sim DHEIE(\alpha,\beta,q)$, then the r^{th} moment of the random variable X can be given by

$$E(X^r) = \sum_{x=1}^{\infty}(x^r - (x-1)^r)S(x),$$

Using the expression of Equation (9), the mean can be given by,

$$E(X) = \sum_{x=1}^{\infty} \frac{\alpha^{\frac{1}{\beta}}\left(1 - q^{\frac{1}{x+1}}\right)}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}, \tag{12}$$

Using the expression of Equation (9), variance can be given by,

$$Var(X) = \sum_{x=1}^{\infty}(2x-1) \frac{\alpha^{\frac{1}{\beta}}\left(1 - q^{\frac{1}{x+1}}\right)}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}} - (E(X))^2. \tag{13}$$

The dispersion index (DsI) is defined as the $variance/|mean|$, it determines whether a given distribution is suitable for equi - dispersed ($DsI = 1$), under-dispersed ($DsI < 1$) or over - dispersed (($DsI > 1$) datasets. DsI is widely used in ecology as a standard measure for measuring repulsion (under dispersion) or clustering (over dispersion).

The numerical values of mean and variance of the DHEIE(α,β,q) distribution for different values of α, β, and q are calculated in Table 1 by using R statistical software. From these values, we compute the dispersion index (DsI), defined as variance/|mean|, to assess the degree of dispersion in the distribution.

- As the parameters α, β, and q increase, the mean, variance, and DsI of the DHEIE distribution increase significantly.
- The variance shows a rapid increase with respect to parameter values, indicating a wide spread in the distribution as parameters grow.
- Dispersion Index (DsI) increases with parameters, reflecting growing overdispersion in the data, especially for higher values of α, β, and q.
- The relationship between the parameters and the statistical measures is non-linear, especially for mid to high values of parameters.
- The DHEIE distribution is highly sensitive to small changes in parameter values; even slight increases in α, β, or q can cause substantial changes in output statistics.
- At low parameter values, especially $(\alpha,\beta) = (0.05, 0.1)$, the mean and variance approach zero, indicating that the distribution is concentrated near zero.

TABLE 1. Mean, Variance, and DsI of DHEIED for various combinations of α, β, and q

α	β	q	Mean	Variance	DsI
	0.1		0.0431	0.0525	1.2181
	0.2		0.0714	0.1501	2.1024
	0.3		0.1126	1.1138	9.8941
↑	0.4	↑	0.1764	6.4705	36.688
0.05	0.5	0.03	0.2742	22.150	80.784
↓	0.6	↓	0.4139	53.369	128.939
	0.7		0.5976	102.629	171.724
	0.8		0.8227	169.725	206.308
	0.9		1.0834	252.662	233.210
0.1			0.1117	0.1588	1.4216
0.2			0.5184	1.4789	2.8530
0.3			1.1091	8.1576	7.3550
0.4	↑	↑	1.9186	39.8566	20.7736
0.5	0.05	0.05	3.0427	150.9535	49.6113
0.6	↓	↓	4.6035	436.6224	94.8458
0.7			6.7220	1016.7723	151.260
0.8			9.4964	2010.2609	211.687
0.9			12.9904	3513.5358	270.472
0.05	0.1		0.0000	0.0000	1.7001
0.10	0.2		0.0026	0.0746	28.2375
0.15	0.3		0.0366	4.1021	112.025
0.20	0.4	↑	0.1579	29.6045	187.438
0.25	0.5	0.5	0.3922	92.3997	235.608
0.30	0.6	↓	0.7185	190.0621	264.542
0.35	0.7		1.0973	309.4172	281.972
0.40	0.8		1.4930	436.9089	292.640
0.45	0.9		1.8805	562.7938	299.277

- The maximum dispersion (highest DsI) occurs when all three parameters are large, suggesting that the DHEIE model is suitable for modeling heavy-tailed or highly variable data.

Therefore, the DHEIE(α, β, q) distribution is highly flexible to model a wide range of dispersion behaviors from repulsion (under-dispersion) to clustering (over-dispersion) through appropriate tuning of its parameters.

3.9. **Probability generating function.** The probability generating function (pgf) of the DHEIE distribution is obtained as,

$$P_X(s) = 1 + \alpha^{\frac{1}{\beta}}(s-1)\sum_{x=1}^{\infty} s^{x-1} \frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}.$$

3.10. **Moment generating function.** The moment generating function (mgf) of the DHEIE is given by,

$$M_X(t) = \sum_{x=0}^{\infty}\sum_{m=0}^{\infty} \frac{(xt)^m}{m!}\left(\alpha^{\frac{1}{\beta}}\left(\frac{1-q^{\frac{1}{x}}}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x}}\right)^{\beta}\right)^{\frac{1}{\beta}}} - \frac{1-q^{\frac{1}{x+1}}}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}\right)\right).$$

The first four partial derivatives of $M_X(t)$, with respect to t at $t = 0$, produce the first four raw moments about the origin. The coefficient of skewness and kurtosis can be computed based on moments. If the pdf given by a random variable X in equation (7) , then the corresponding r^{th} moments can be computed, the behavior of mean, variance, skewness and kurtosis for selected values of α, β, q using R software when the upper limit is finite.

3.11. **Shannon entropy.** Shannon entropy is a way to measure how unpredictable or uncertain a situation is when dealing with probabilities. If you have a random variable X that can take on several values with certain probabilities, the Shannon entropy tells you how much "surprise" is involved. The formula looks like this:

$$S_H(X) = -\sum_{x} p(x)\log p(x).$$

Here, $p(x)$ represents the probability that X takes the value x. The more evenly spread the probabilities are, the higher the entropy, meaning more uncertainty and less information.

$$S(X) = -\sum_{x=0}^{\infty}\left[\frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x}}\right)^{\beta}\right)^{\frac{1}{\beta}}} - \frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}\right]$$
$$\times \log\left[\frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x}}\right)^{\beta}\right)^{\frac{1}{\beta}}} - \frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}\right].$$

The numerical illustration of Shannon entropy for various parameter combinations is presented in Table 2. From this we can infer the following points.

Table 2. Entropy of DHEIED for various combinations of α, β, and q.

α	β	q = 0.02	q = 0.1	q = 0.9
	0.1	2.415336	1.762221	0.084393
	0.2	2.477083	1.825203	0.112271
↑	0.3	2.532580	1.882129	0.138394
0.5	0.4	2.582587	1.933579	0.162281
↓	0.5	2.627805	1.980160	0.183834
	0.6	2.668832	2.022428	0.203142
	0.7	2.706172	2.060876	0.220379
	0.8	2.740256	2.095936	0.235755
0.1		1.186696	0.770590	0.015464
0.2		1.730851	1.169968	0.047773
0.3	↑	2.106373	1.488284	0.088846
0.4	0.5	2.393948	1.753656	0.134920
0.6	↓	2.825191	2.177171	0.234256
0.7		2.996012	2.351164	0.285334
0.8		3.146500	2.506733	0.336512

- Shannon entropy increases as either α or β increases, indicating that both parameters contribute positively to the uncertainty or disorder in the system.
- For any fixed pair of α and β, the entropy decreases as the entropic index q increases. This reflects the property that higher q values in Tsallis or Rényi like frameworks give more weight to the dominant probabilities, thus lowering the entropy.

- The rate of increase in entropy with respect to α or β is more pronounced when q is small, suggesting that lower q values are more sensitive to parameter changes.
- When either α or β is small, entropy values tend to be low, especially for larger q, implying reduced uncertainty in more concentrated or less spread-out distributions.
- The entropy surface (over α, β, and q) appears smooth and monotonic, with no irregular or non-monotonic behavior, which is consistent with well behaved parametric models.
- In summary, both parameters α and β enhance entropy, while increasing q suppresses it, indicating a clear and consistent interaction between shape/scale parameters and the entropy index.

3.12. **Order statistics.** Order statistics (OS) play a significant role in various areas of statistical theory and practice. Consider a random sample $X_1, X_2, \ldots, X_n$ drawn from the DHEIE(α, β, q) model, and let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ represent the corresponding order statistics. The cdf for the i^th order statistic $X_{i:n}$, for an integer value of x, can be expressed as follows:

$$\begin{aligned}
F_{i:n}(x;\alpha,\beta,q) &= \sum_{k=i}^{n}\binom{n}{k}[F(x;\alpha,\beta,q)]^{k}[1-F(x;\alpha,\beta,q)]^{n-k},\\
&= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\Delta_j(n,k)[F(x;\theta,\alpha,\beta)]^{k+j},\\
&= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\Delta_j(n,k)\left[1-\frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}\right]^{k+j},
\end{aligned}$$

where $\Delta_j(n,k) = (-1)^j\binom{n}{k}\binom{n-k}{j}$.

The pmf of the k^{th} order statistic is given by:

$$f_X(x) = F_X(x) - F_X(x-1).$$

$$\begin{aligned}
f_X(x) &= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\Delta_j(n,k)\left[\left(1-\frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x+1}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}}\right)^{k+j}-\left(1-\frac{\alpha^{\frac{1}{\beta}}\left(1-q^{\frac{1}{x}}\right)}{\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x}}\right)^{\beta}\right)^{\frac{1}{\beta}}}\right)^{k+j}\right],\\
&= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\Delta_j(n,k)\left[\left(1+\alpha^{\frac{1}{\beta}}\left(q^{\frac{1}{x+1}}-1\right)\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)^{-\frac{1}{\beta}}\right)^{j+k}\right.\\
&\qquad \left.-\left(1+\alpha^{\frac{1}{\beta}}\left(q^{\frac{1}{x}}-1\right)\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x}}\right)^{\beta}\right)^{-\frac{1}{\beta}}\right)^{j+k}\right].
\end{aligned}$$

3.13. **Stress-strength parameter.** The stress-strength parameter R serves as a key metric for assessing component reliability. It is defined in the context of a random variable Y, representing the strength of a component, subjected to a random stress X. The estimation of R has been extensively explored in statistical literature, particularly under the assumption that X and Y are independent and identically distributed (i.i.d.). Numerous researchers have contributed to this area of study, with extensive discussions available in the literature. it has been widely applicable in diverse fields such as medicine, engineering, and psychology. In the discrete case, the stress-strength model is defined as

$$R = P(Y > X) = \sum_{x=0}^{\infty} p_Y(x) F_X(x),$$

where $p_Y(x)$ is the pmf of Y and $F_X(x)$ is the cdf of X.

Let $Y \sim \text{DHEIE}(\boldsymbol{\gamma}_1)$ and $X \sim \text{DHEIE}(\boldsymbol{\gamma}_2)$, where $\boldsymbol{\gamma}_1 = (\alpha_1, \beta_1, q_1)^T$ and $\boldsymbol{\gamma}_2 = (\alpha_2, \beta_2, q_2)^T$. Then, using equations (7) and (8), we have

$$R = \sum_{x=0}^{\infty} \left((q_1^{\frac{1}{1+x}} - 1) \left(1 - \bar{\alpha}_1 (1 - q_1^{\frac{1}{1+x}})^{\beta_1}\right)^{-\frac{1}{\beta_1}} - (q_1^{\frac{1}{x}} - 1) \left(1 - \bar{\alpha}_1 (1 - q_1^{\frac{1}{x}})^{\beta_1}\right)^{-\frac{1}{\beta_1}} \right)$$
$$\times \alpha_1^{\frac{1}{\beta_1}} \left(1 + \alpha_2^{\frac{1}{\beta_2}} (q_2^{\frac{1}{1+x}} - 1) \left(1 - \bar{\alpha}_2 (1 - q_2^{\frac{1}{1+x}})^{\beta_2}\right)^{-\frac{1}{\beta_2}} \right).$$

The stress-strength reliability parameter for different parameter values is numerically computed and presented in Table 3. From this, we can infer that the reliability (R value) improves significantly when the strength distribution dominates the stress distribution, particularly when $q_2 > q_1$.

4. Estimation

In the following section, we employ five estimation methods namely, maximum likelihood (ML), Ordinary least squares (OLS), Weighted least squares (WLS), Cramér–von Mises (CM), and Anderson–Darling (AD) to estimate the parameters of the proposed distribution.

4.1. **Maximum likelihood estimation.** Let $X_1, X_2, \ldots, X_n$ be a random sample from the DHEIE distribution with parameters α, β, q. The log-likelihood function is given by:

$$\ell(\alpha, \beta, q) = \sum_{i=1}^{n} \log\left[p(x_i; \alpha, \beta, q)\right].$$

Which is obtained as,

$$\ell(\alpha, \beta, q) = \frac{n}{\beta} \log \alpha + \sum_{i=1}^{n} \log \left[\frac{\left(1 - q^{\frac{1}{x_i}}\right)}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x_i}}\right)^{\beta}\right)^{\frac{1}{\beta}}} - \frac{\left(1 - q^{\frac{1}{x_i+1}}\right)}{\left(1 - \bar{\alpha}\left(1 - q^{\frac{1}{x_i+1}}\right)^{\beta}\right)^{\frac{1}{\beta}}} \right].$$

Table 3. Combined Table of R Values with Grouped q_1 and q_2

α_1	β_1	α_2	β_2	**R Value**
$q_1 = 0.5$, $q_2 = 0.9$				
0.2	0.1	0.4	0.3	0.04592
0.4	0.5	0.8	0.6	0.22837
0.6	1.0	1.2	0.9	0.34653
0.8	1.5	1.6	1.2	0.41322
$q_1 = 0.3$, $q_2 = 0.6$				
α_1	β_1	α_2	β_2	**R Value**
0.2	0.1	0.4	0.3	0.16466
0.4	0.3	0.8	0.6	0.32764
0.6	1.0	1.2	0.9	0.41980
0.8	1.5	1.6	1.2	0.45779
$q_1 = 0.9$, $q_2 = 0.5$				
α_1	β_1	α_2	β_2	**R Value**
0.2	0.1	0.4	0.3	0.00019
0.4	0.3	0.8	0.6	0.01049
0.6	1.0	1.2	0.9	0.03906
0.8	1.5	1.6	1.2	0.05015

Hence, the likelihood equations of α, β, q is obtained by solving the following non-linear equations,

$$\frac{\partial \ell}{\partial q} = \sum_{i=1}^{n} \frac{1}{\frac{A_i}{D_i} - \frac{A_{i+1}}{D_{i+1}}} \left[\frac{D_i \frac{\partial A_i}{\partial q} - A_i \frac{\partial D_i}{\partial q}}{D_i^2} - \frac{D_{i+1} \frac{\partial A_{i+1}}{\partial q} - A_{i+1} \frac{\partial D_{i+1}}{\partial q}}{D_{i+1}^2} \right], \tag{14}$$

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\beta \alpha} - \sum_{i=1}^{n} \frac{\left[\frac{A_i^{\beta+1}}{\beta D_i^2} \left(1 - \bar{\alpha} A_i^{\beta}\right)^{\frac{1}{\beta}-1} - \frac{A_{i+1}^{\beta+1}}{\beta D_{i+1}^2} \left(1 - \bar{\alpha} A_{i+1}^{\beta}\right)^{\frac{1}{\beta}-1} \right]}{\left(\frac{A_i}{D_i} - \frac{A_{i+1}}{D_{i+1}} \right)}, \tag{15}$$

$$\frac{\partial \ell}{\partial \beta} = -\frac{n}{\beta^2} \log \alpha + \sum_{i=1}^{n} \frac{1}{\frac{A_i}{D_i} - \frac{A_{i+1}}{D_{i+1}}} \left[-\frac{A_i}{D_i^2} \frac{\partial D_i}{\partial \beta} + \frac{A_{i+1}}{D_{i+1}^2} \frac{\partial D_{i+1}}{\partial \beta} \right]. \tag{16}$$

Where,

$$A_i = 1 - q^{\frac{1}{x_i}}, \quad A_{i+1} = 1 - q^{\frac{1}{x_i+1}},$$

$$D_i = \left(1 - \bar{\alpha} A_i^{\beta}\right)^{1/\beta}, \quad D_{i+1} = \left(1 - \bar{\alpha} A_{i+1}^{\beta}\right)^{1/\beta}.$$

The maximum likelihood estimators (MLEs) of $\Theta = (\alpha, \beta, q)^T$, denoted by $\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{q})^T$, can be obtained by numerically solving the likelihood equations (14)– (14)–(20). A suitable approach for this is a multivariable optimization method, such as the four-parameter Newton-Raphson algorithm.

4.2. **Ordinary least square and Weighted least square estimation.** The method of Least square estimation was proposed by [18] to determine the unknown parameters by minimizing the distance between uniformized order statistics vector and the corresponding vector of expected values. Given $X_1, X_2, X_3, \ldots, X_n$, as the random sample of size n, taken from a continuous distribution function, then let $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq \ldots \leq X_{(n)}$ be its corresponding order statistics. Then the expected value and variance of the empirical cumulative distribution function (ecdf) are defined as follows,

$$E\big(F(X_{(i)}\big) = \frac{i}{n+1}, \qquad i = 1, 2, \ldots, n \tag{17}$$

$$Var\big(F(X_{(i)}\big) = \frac{i(n-i+1)}{(n+1)^2(n+2)}, \qquad i = 1, 2, \ldots, n. \tag{18}$$

From the above equation, we can derive the OLS and the WLS estimators of the unknown parameters α, β and q.

The OLSE and WLSE method for estimating the unknown parameters, (α, β, q) of the DHEIE can be obtained by minimising the following equation,

$$S(\alpha, \beta, q) = \sum_{i=1}^{n} W_i \left(F(X_{(i)}; \alpha, \beta, q) - \frac{i}{n+1} \right) \Bigg)^2, \tag{19}$$

where $W_i = 1$ for all $i = 1, 2, \ldots, n$ in the case of OLSE method which results in estimators $\hat{\alpha}_{OLSE}$, $\hat{\beta}_{OLSE}$ and $\hat{q}_{OLSE}$ and $W_i = \frac{1}{Var\left(F(X_{(i)}\right)}$ in Equation (19) in the WLSE method that produce estimators $\hat{\alpha}_{WLSE}$, $\hat{\beta}_{WLSE}$ and $\hat{q}_{WLSE}$. Further the estimation of parameters α, β and q of the DHEIE is obtained by solving the following non-linear equations.

$$\frac{\partial S}{\partial \alpha} = 2 \sum_{i=1}^{n} W_i \left(F(X_{(i)}; \alpha, \beta, q) - \frac{i}{n+1} \right) \frac{\partial F(X_{(i)}; \alpha, \beta, q)}{\partial \alpha} = 0,$$

$$\frac{\partial S}{\partial \beta} = 2 \sum_{i=1}^{n} W_i \left(F(X_{(i)}; \alpha, \beta, q) - \frac{i}{n+1} \right) \frac{\partial F(X_{(i)}; \alpha, \beta, q)}{\partial \beta} = 0,$$

$$\frac{\partial S}{\partial q} = 2 \sum_{i=1}^{n} W_i \left(F(X_{(i)}; \alpha, \beta, q) - \frac{i}{n+1} \right) \frac{\partial F(X_{(i)}; \alpha, \beta, q)}{\partial q} = 0.$$

where,

$$\frac{\partial F(X_{(i)}; \alpha, \beta, q)}{\partial \alpha} = \frac{\alpha^{-1+1/\beta} \left[1 - \bar{\alpha} \left(1 - q^{\frac{1}{x+1}} \right)^{\beta} \right]^{-1/\beta} \left(1 - q^{\frac{1}{x+1}} \right)}{\beta}, \tag{20}$$

$$\frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial \beta} = -\frac{\alpha^{1/\beta}\left[1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right]^{-1/\beta}\left(1-q^{\frac{1}{x+1}}\right)}{\beta^2} \times \left(\log(\alpha)-\log\left(1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right)\right), \tag{21}$$

$$\frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial q} = \frac{\alpha^{1/\beta}\left[1-\bar{\alpha}\left(1-q^{\frac{1}{x+1}}\right)^{\beta}\right]^{-1/\beta} q^{-1+\frac{1}{x+1}}}{x+1}. \tag{22}$$

4.3. **Anderson-Darling estimation.** The Anderson-Darling estimation (ADE) is based on Anderson-Darling statistic proposed by [19]. The estimator $q_{\hat{ADE}}$, $\alpha_{\hat{ADE}}$ and $\beta_{\hat{ADE}}$, can be obtained by minimizing following equation with respect to α, β, q. parameter.

$$A(\alpha,\beta,q) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\Big(\log(F(X_{(i)})) + \log(1-F(X_{(n+1-i)}))\Big). \tag{23}$$

Furthermore, these estimation of unknown parameters are done by solving the following non-linear equations,

$$\frac{\partial A(\alpha,\beta,q)}{\partial \alpha} = \sum_{i=1}^{n}(2i-1)\left[\frac{\phi_{1i}}{F(X_{(i)})} - \frac{\phi_{1(n+1-i)}}{1-F(X_{(n+1-i)})}\right] = 0,$$
$$\frac{\partial A(\alpha,\beta,q)}{\partial q} = \sum_{i=1}^{n}(2i-1)\left[\frac{\phi_{2i}}{F(X_{(i)})} - \frac{\phi_{2(n+1-i)}}{1-F(X_{(n+1-i)})}\right] = 0,$$
$$\frac{\partial A(\alpha,\beta,q)}{\partial \beta} = \sum_{i=1}^{n}(2i-1)\left[\frac{\phi_{3i}}{F(X_{(i)})} - \frac{\phi_{2(n+1-i)}}{1-F(X_{(n+1-i)})}\right] = 0,$$

where

$$\phi_{1i} = \frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial \alpha}, \phi_{2i} = \frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial \beta}, \phi_{3i} = \frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial q},$$

has similiar expression as (20)-(22).

4.4. **Cramér-von Mises estimation.** Similar to ADE method the Cramér-von Mises estimator (CME) is a type of minimum distance estimator, also known as maximum goodness-of fit estimators. This estimator is based on the difference between the estimate of cdf and the ecdf. The CME of $\alpha_{\hat{CME}}$, $\beta_{\hat{CME}}$ and $q_{\hat{CME}}$, are resulted from minimizing the equation in (24) with respect to α, β and q.

$$C(\alpha,\beta,q) = \frac{1}{12n} + \sum_{i=1}^{n}\left(F(X_{(i)}) - \frac{2i-1}{2n}\right)^2 \tag{24}$$

i.e. the unknown parameters are obtained by solving the following non-linear equations

$$\frac{\partial C}{\partial \alpha} = 2\sum_{i=1}^{n}\left(F(X_{(i)};\alpha,\beta,q) - \frac{2i-1}{2n}\right)\frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial \alpha} = 0, \tag{25}$$

$$\frac{\partial C}{\partial \beta} = 2\sum_{i=1}^{n}\left(F(X_{(i)};\alpha,\beta,q) - \frac{2i-1}{2n}\right)\frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial \beta} = 0, \tag{26}$$

$$\frac{\partial C}{\partial q} = 2\sum_{i=1}^{n}\left(F(X_{(i)};\alpha,\beta,q) - \frac{2i-1}{2n}\right)\frac{\partial F(X_{(i)};\alpha,\beta,q)}{\partial q} = 0. \tag{27}$$

where the partial derivatives of the equation (25)-(27) are obtained in (20)-(22).

5. Simulation

In this section, we will go over simulation studies for various parameter combinations in order to assess the performance of the five different parameter estimation methods, including MLE, OLSE, WLSE, ADE and CME. We generated 1000 simulations for different choices of size, n ($n = 20, 100, 500$). In each simulated data a n observations/sample data points are generated from the DHEIE distribution for different combination of parameters such as $(\alpha, \beta, q) = \{(0.5, 0.5, 0.5), (1.5, 1, 0.8)\}$. Each sample observations are generated from the inverse cdf of DHEIE distribution. Using the five estimation methods, we estimated parameter values of each simulated data by applying numerical algorithm to solve the non-linear equations given for all the five methods.

The bias and mean squared error (MSE) were computed and depicted in Table 4 and 5 to analyse the performance of the five methods, simulated for different sample sizes(n) and for various choices of parameters. Table 4 and 5 shows that all the estimation techniques perform effectively for various parameter selections, and the bias and MSE decrease as the sample size increases.

6. Applications

This section discusses the advantages of the newly proposed DHEIE distribution over some commonly used discrete distributions. The performance of the DHEIE distribution is compared with the following competitive distributions such as the discrete Burr XII distribution (DBXII), the discrete Bilal distribution (DB), the discrete Burr-Hatke distribution (DBH), the discrete exponentiated Rayleigh distribution (DER), the discrete length-biased exponential distribution (DLBE), the discrete Pareto distribution (DPr) and the Poisson distribution (P).

Table 4. Bias and MSE of (α, β, q) for various parameter values obtained using five estimation methods with different sample sizes.

Parameters			Bias			MSE		
Setting	n	Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{q}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{q}$
$\alpha = 0.5, \beta = 0.5, q = 0.5$	20	MLE	0.040	0.048	0.421	0.024	0.036	0.0410
		CME	0.032	0.027	0.039	0.034	0.067	0.074
		ADE	-0.007	-0.017	0.022	0.022	0.050	0.063
		OLSE	0.044	-0.045	-0.043	0.037	0.074	0.069
		WLSE	0.046	-0.044	-0.051	0.040	0.076	0.063
	100	MLE	0.006	0.002	0.003	0.004	0.011	0.014
		CME	0.004	0.003	0.006	0.006	0.013	0.017
		ADE	-0.003	-0.005	0.004	0.005	0.011	0.013
		OLSE	0.006	-0.012	-0.003	0.006	0.014	0.012
		WLSE	0.007	-0.011	-0.012	0.006	0.014	0.021
	500	MLE	0.003	0.002	0.002	0.001	0.002	0.003
		CME	0.003	0.003	0.002	0.001	0.003	0.002
		ADE	0.001	0.001	0.001	0.001	0.003	0.005
		OLSE	0.004	0.001	0.001	0.003	0.004	0.002
		WLSE	0.004	0.001	0.002	0.001	0.003	0.004

Data set I: Deaths due to coronavirus in China. The first data set consists of the daily number of deaths due to coronavirus (COVID-19) in China from 23 January to 28 March 2020. The data were collected from the website https://www.worldometers.info/coronavirus/country/china/. The recorded number of daily deaths is as follows:

Data: 8, 16, 15, 24, 26, 26, 38, 43, 46, 45, 57, 64, 65, 73, 73, 86, 89, 97, 108, 97, 146, 121, 143, 142, 105, 98, 136, 114, 118, 109, 97, 150, 71, 52, 29, 44, 47, 35, 42, 31, 38, 31, 30, 28, 27, 22, 17, 22, 11, 7, 13, 10, 14, 13, 11, 8, 3, 7, 6, 9, 7, 4, 6, 5, 3, 5.

The MLEs, along with their corresponding standard errors and goodness of fit measures for the fitted models, are presented in Table 6. From Table 6, it is evident that the DHEIE distribution yields the lowest values of AIC, BIC, and KS statistic, along with the highest p-value. These results indicate that the DHEIE distribution provides a better fit compared to the other competing distributions. The fitted pdf and cdf plots of the DHEIE distribution for COVID-19 deaths in China are shown in Figure 3.

Data set II: Daily deaths due to coronavirus in Pakistan. The second data set consists of the daily number of deaths due to coronavirus (COVID-19) in Pakistan from 18 March to 30 June 2020. The

Table 5. Bias and MSE of (α, β, q) for various parameter values obtained using five estimation methods with different sample sizes

Parameters			Bias			MSE		
Setting	n	Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{q}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{q}$
$\alpha = 1.5, \beta = 1.0, q = 0.8$	20	MLE	0.101	0.084	0.071	0.086	0.061	0.055
		CME	0.085	0.052	0.049	0.064	0.049	0.031
		ADE	-0.016	-0.028	0.031	0.038	0.020	0.023
		OLS	0.096	-0.068	-0.052	0.097	0.084	0.071
		WLS	0.121	-0.091	-0.065	0.096	0.092	0.087
	100	MLE	0.019	0.017	0.009	0.028	0.017	0.014
		CME	0.009	0.006	0.008	0.011	0.019	0.022
		ADE	-0.008	-0.011	0.012	0.009	0.018	0.017
		OLS	0.011	-0.019	-0.011	0.008	0.018	0.019
		WLS	0.017	-0.023	-0.018	0.009	0.018	0.019
	500	MLE	0.007	0.004	0.004	0.005	0.004	0.004
		CME	0.004	0.005	0.003	0.003	0.003	0.002
		ADE	0.002	0.003	0.002	0.003	0.005	0.004
		OLS	0.006	0.004	0.004	0.008	0.006	0.005
		WLS	0.008	0.006	0.004	0.003	0.004	0.004

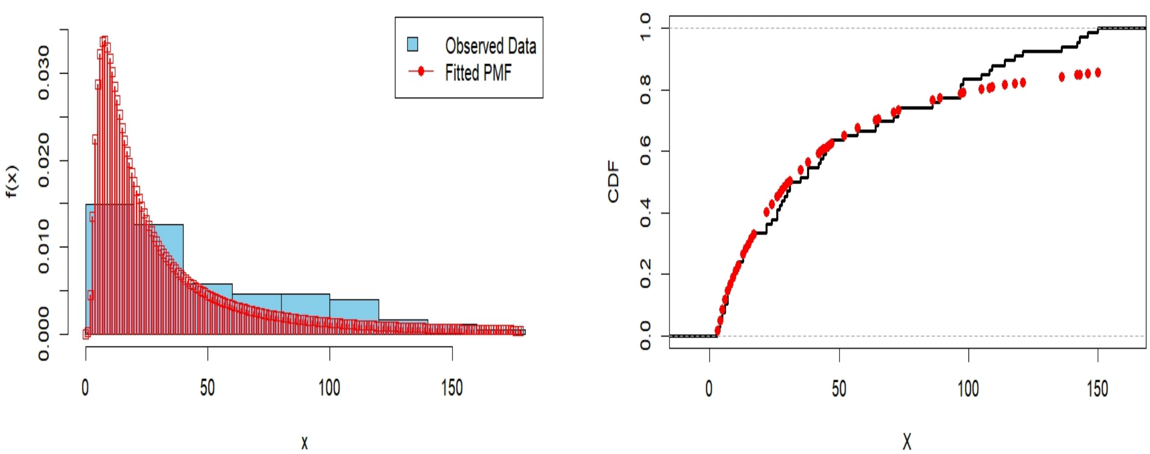

Figure 3. Fitted pdf and cdf plots of COVID-19 deaths in China.

TABLE 6. MLEs, standard errors, and model comparison criteria for COVID-19 deaths in China

Model	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$-\log L$	AIC	BIC	KS (p-value)
DHEIED	18.8417	0.0398	28.291	328.63	663.26	669.83	0.1290 (0.2104)
	(9.751)	(0.067)	(6.448)				
DBXII	0.9788	6.3999	–	374.49	752.99	757.38	0.3607 (0.0000)
	(0.0389)	(17.562)					
DER	34.054	0.5246	–	347.23	698.45	702.83	0.2932 (0.0000)
	(169.93)	(5.2344)					
DPr	0.2863	–	–	379.07	760.14	762.33	0.3816 (0.0000)
	(0.0352)						
DLBE	25.122	–	–	330.52	663.03	665.22	0.1718 (0.0407)
	(2.1866)						
DB	0.9834	–	–	330.07	662.14	664.33	0.1655 (0.0538)
	(12.114)						
DBH	0.9998	–	–	461.02	924.04	926.23	0.8119 (0.0000)
	(0.0019)						
P	49.737	–	–	1409.8	2821.6	2823.7	0.4975 (0.0000)
	(0.8681)						

data were collected from the website https://www.worldometers.info/coronavirus/country/Pakistan. The recorded number of daily deaths is as follows:

Data: 1, 6, 6, 4, 4, 4, 1, 20, 5, 2, 3, 15, 17, 7, 8, 25, 8, 25, 11, 25, 16, 16, 12, 11, 20, 31, 42, 32, 23, 17, 19, 38, 50, 21, 14, 37, 23, 47, 31, 24, 9, 64, 39, 30, 36, 46, 32, 50, 34, 32, 34, 30, 28, 35, 57, 78, 88, 60, 78, 67, 82, 68, 97, 67, 65, 105, 83, 101, 107, 88, 178, 110, 136, 118, 136, 153, 119, 89, 105, 60, 148, 59, 73, 83, 49, 137, 91.

The MLEs, corresponding standard errors, and model comparison metrics are reported in Table 7.

From Table 7, it is evident that the DHEIE distribution yields the lowest values of AIC, BIC, and KS statistic, along with the highest p-value. These results indicate that the DHEIE distribution provides a better fit compared to the other competing distributions. The fitted pdf and cdf plots of the DHEIE distribution for COVID-19 deaths in Pakistan are shown in Figure 4.

Table 7. MLEs, standard errors, and model comparison criteria for COVID-19 deaths in Pakistan

Model	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$-\log L$	AIC	BIC	KS (p-value)
DHEIED	5.1674	0.9999	16.0114	451.76	909.53	916.93	0.1826 (0.2426)
	(7.7459)	(0.01001)	(0.0001)				
DBXII	0.9816	15.499	–	497.12	998.24	1003.2	0.3500 (0.0000)
	(0.0227)	(19.249)					
DER	33.716	0.5293	–	452.54	909.09	914.02	0.2473 (0.0000)
	(228.83)	(7.1851)					
DPr	0.2834	–	–	503.61	1009.2	1011.6	0.3556 (0.0000)
	(0.03038)						
DBH	0.9997	–	–	613.80	1229.6	1232.1	0.7876 (0.0000)
	(0.0016)						
DP	50.057	–	–	1713.0	3428.1	3430.5	0.4579 (0.0000)
	(1713.0)						

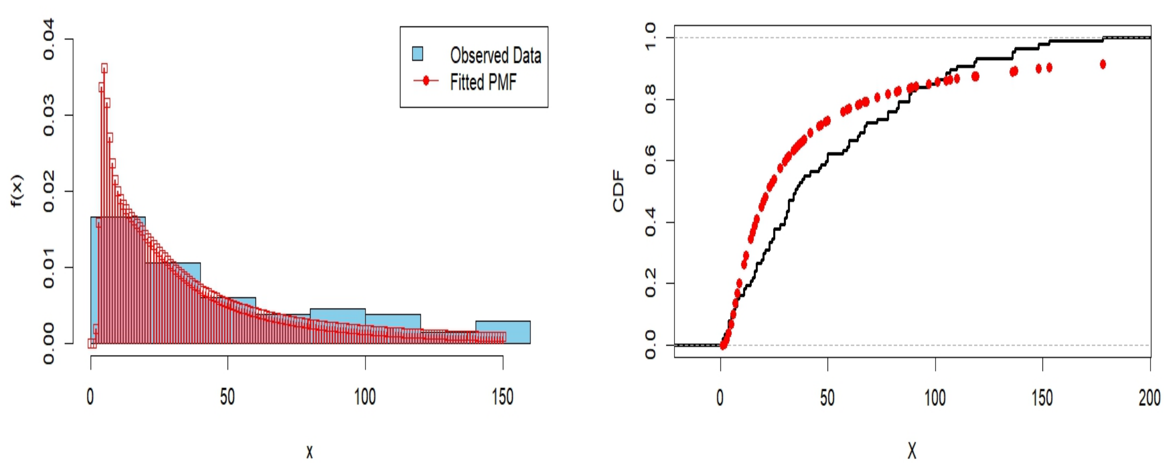

Figure 4. Fitted pdf and cdf plots of COVID-19 deaths in Pakistan.

7. Conclusion

In this paper, we introduce a novel discrete distribution called the Harris extended inverse exponential distribution. We examine several key structural properties of the proposed distribution and provide a detailed analysis of its mean and variance, supported by numerical illustrations. The model parameters are estimated using various methods, including MLE, ADE, CME, OLSE, and WLSE. To evaluate the performance of these estimation techniques, we conduct a comprehensive simulation study. In addition, we demonstrate the practical relevance and flexibility of the proposed distribution by analyzing two real-world data sets related to COVID-19 from China and Pakistan. This new model can serve as a viable alternative to the existing discrete distributions in the statistical literature.

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

References

[1] T. Nakagawa, S. Osaki, The discrete Weibull distribution, IEEE Trans. Reliab. 24 (1975) 300-301. https://doi.org/10.1109/tr.1975.5214915.

[2] W.E. Stein, R. Dattero, A new discrete Weibull distribution, IEEE Trans. Reliab. 33 (1984) 196-197. https://doi.org/10.1109/tr.1984.5221777

[3] W.J. Padgett, J.D. Spurrier, On discrete failure models, IEEE Trans. Reliab. 34 (1985) 253-256. https://doi.org/10.1109/tr.1985.5222137.

[4] H. Sato, M. Ikota, A. Sugimoto, H. Masuda, A new defect distribution metrology with a consistent discrete exponential formula and its applications, IEEE Trans. Semicond. Manuf. 12 (1999) 409-418. https://doi.org/10.1109/66.806118.

[5] D. Roy, The discrete normal distribution, Commun. Stat. Theory Methods 32 (2003) 1871-1883. https://doi.org/10.1081/sta-120023256.

[6] D. Roy, Discrete Rayleigh distribution, IEEE Trans. Reliab. 53 (2004) 255-260. https://doi.org/10.1109/tr.2004.829161.

[7] S. Inusah, T.J. Kozubowski, A discrete analogue of the Laplace distribution, J. Stat. Plann. Inference 136 (2006) 1090-1102. https://doi.org/10.1016/j.jspi.2004.08.014.

[8] T.J. Kozubowski, S. Inusah, A skew Laplace distribution on integers, Ann. Inst. Stat. Math. 58 (2006) 555-571. https://doi.org/10.1007/s10463-005-0029-1.

[9] H. Krishna, P.S. Pundir, Discrete Burr and discrete Pareto distributions, Stat. Methodol. 6 (2009) 177-188. https://doi.org/10.1016/j.stamet.2008.07.001.

[10] M.A. Jazi, C.-D. Lai, M.H. Alamatsaz, A discrete inverse Weibull distribution and estimation of its parameters, Stat. Methodol. 7 (2010) 121-132. https://doi.org/10.1016/j.stamet.2009.11.001.

[11] C. Bracquemond, O. Gaudoin, A survey on discrete lifetime distributions, Int. J. Reliab. Qual. Saf. Eng. 10 (2003) 69-98. https://doi.org/10.1142/s0218539303001007.

[12] S. Chakraborty, S.H. Ong, Mittag-Leffler function distribution—a new generalization of hyper-Poisson distribution, J. Stat. Distrib. Appl. 4 (2017) 1-17. https://doi.org/10.1186/s40488-017-0060-9.

[13] E.-E.A.A. Aly, L. Benkherouf, A new family of distributions based on probability generating functions, Sankhya B 73 (2011) 70-80. https://doi.org/10.1007/s13571-011-0017-9.

[14] G.M. Cordeiro, M. Mansoor, S.B. Provost, The Harris extended Lindley distribution for modeling hydrological data, Chilean J. Stat. 10 (2019).

[15] L. Tomy, G. Veena, C. Chesneau, A new three-parameter lifetime distribution for environmental data analysis: the Harris extended modified Lindley distribution, Int. J. Data Sci. Anal. (2024) 1-16. https://doi.org/10.1007/s41060-024-00505-0.

[16] J. Bengalath, B. Punathumparambath, Harris extended inverted Kumaraswamy distribution: Properties and applications to COVID-19 data, Int. J. Data Sci. Anal. (2024) 1-23. https://doi.org/10.1007/s41060-024-00639-1.

[17] S.P. Thomas, L. Tomy, K.K. Jose, Discrete Harris extended Weibull distribution and applications, Stat. Appl. (2023) 63.

[18] J. Swain, S. Venkatraman, J. Wilson, Least-squares estimation of distribution functions in Johnson's translation system, J. Stat. Comput. Simul. 29 (1988) 271-297.https://doi.org/10.1080/00949658808811068.

[19] T.W. Anderson, D.A. Darling, Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, Ann. Math. Stat. 23 (1952) 193-212.https://doi.org/10.1214/aoms/1177729437.

[20] J. Bengalath, B. Punathumparambath, Harris extended Dagum distribution: properties and applications to reliability analysis, Life Cycle Reliab. Saf. Eng. (2025) 1-18. https://doi.org/10.1007/s41872-025-00344-3.