AIMS Mathematics, 11(8): 26791–26822.
DOI: 10.3934/math.20261075
Received: 27 May 2026
Revised: 28 July 2026
Accepted: 19 August 2026
Published: 25 August 2026

https://www.aimspress.com/journal/Math

Research article

A new shifted Lagrangian exponential Poisson model with applications to overdispersed and zero-inflated count data

Fadal Abdullah Ali Aldhufairi*

Department of Mathematics, King Khalid University, Abha 61421, KSA

* **Correspondence:** Email: faldhufairi@kku.edu.sa.

Abstract: This paper introduces a new class of discrete distributions, named the shifted Lagrangian exponential Poisson (SLEP) distribution, constructed via a Lagrangian functional equation for the probability-generating function. The resulting model admits an explicit probability mass function derived using Lagrange inversion and is supported on the positive integers. Its structure provides a framework for modeling overdispersed count data, with parameters controlling dispersion and tail behavior. A zero-inflated extension of the SLEP distribution is proposed to model extra zeros, based on a mixture representation of a point mass at zero and the SLEP distribution for positive counts. This formulation enables the simultaneous modeling of zero inflation and overdispersion within a unified framework. A regression extension is further proposed to incorporate covariate effects through log and logit link functions for the positive-count and zero-inflation components, respectively. The model is examined on the DoctorVisits dataset. The empirical results indicate that the zero-inflated SLEP model performs better than the traditional alternatives in log-likelihood, Akaike information criterion, Bayesian information criterion, and chi-square goodness-of-fit measures, including Poisson, negative binomial, and zero-inflated Poisson models. The bootstrap analysis indicates that the estimation procedure is stable, with insignificant bias, accurate standard errors, and satisfactory coverage probabilities. The proposed framework provides a theoretically robust and adaptable solution for modeling intricate count data characterized by overdispersion and zero inflation, and is applicable in medical, actuarial, and social science contexts.

Keywords: count distributions; Lagrangian distributions; overdispersion; zero inflation; maximum likelihood estimation; count regression
Mathematics Subject Classification: 62E15, 62F10

1. Introduction

Count data arise naturally in many scientific disciplines, including epidemiology, public health, actuarial science, reliability analysis, ecology, insurance, and social sciences. Such data commonly represent the number of occurrences of an event within a fixed period or region, for example, hospital visits, insurance claims, disease incidences, accidents, or equipment failures. Statistical modeling of count responses, therefore, plays a fundamental role in modern applied statistics and data analysis [1–3].

The Poisson distribution remains the classical starting point for count data analysis because of its mathematical simplicity and tractable probabilistic structure. Its probability mass function (PMF) is given by

$$P_{\text{Poisson}}(Y = y; \lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!}, \qquad y = 0, 1, 2, \ldots,$$

where $\lambda > 0$ denotes both the mean and variance of the distribution [4]. In practice, however, the assumption that the mean equals the variance is often violated. Real count data often exhibit overdispersion, underdispersion, and excess zeros arising from structural or sampling mechanisms [5–7].

Several extensions of the Poisson model have therefore been developed to overcome these limitations. For example, the negative binomial (NB) distribution accommodates overdispersion through an additional dispersion parameter and has the PMF

$$P_{\text{NB}}(Y = y; \mu, k) = \frac{\Gamma(y + k)}{\Gamma(k)y!}\left(\frac{k}{\mu + k}\right)^{k}\left(\frac{\mu}{\mu + k}\right)^{y},$$

where μ is the mean, k is the dispersion parameter, and $\Gamma(\cdot)$ denotes the Gamma function [8]. Zero-inflated models provide another approach to handling structural zeros. A popular example is the zero-inflated Poisson (ZIP) model, defined by

$$P_{\text{ZIP}}(Y = y; \pi, \lambda) = \pi\mathbb{I}(y = 0) + (1 - \pi)P_{\text{Poisson}}(Y = y; \lambda),$$

where $0 \leq \pi \leq 1$ is the probability of structural zeros [9, 10]. Other extensions, including generalized Poisson and related count distributions, have also been proposed to accommodate a wider range of dispersion patterns and count data structures [11, 12].

Furthermore, several studies have emphasized the importance of accounting for overdispersion in real-world count-data applications, since ignoring extra-Poisson variation may adversely affect the efficiency of parameter estimation and statistical inference [13].

Zero-inflated models have received substantial attention because many biomedical, health, ecological, and insurance datasets contain more zero observations than expected under standard

count distributions. These approaches typically combine a degenerate distribution at zero with a baseline count distribution such as the Poisson or NB distribution [7, 9, 14]. Although such mixture-based constructions provide greater modeling flexibility, their estimation and interpretation may become more challenging because the observed counts are generated through multiple latent stochastic mechanisms.

An alternative framework for constructing flexible count distributions is provided by the Lagrangian probability models introduced by Consul and Shenton [15] and further developed by Consul and Famoye [16]. These distributions are generated through functional equations involving probability-generating functions (PGFs) together with the Lagrange expansion theorem. The Lagrangian framework has produced several useful count models capable of accommodating diverse dispersion structures while preserving tractable analytical forms [17, 18].

More recently, flexible count distributions have been extended to accommodate increasingly complex count data structures and varying dispersion behavior. For example, [19] proposed a rational power-shifted Lagrangian distribution capable of accommodating flexible dispersion characteristics, and [20] demonstrated the practical usefulness of alternative count distributions in applied data settings.

Despite these attractive properties, relatively limited attention has been paid to the development of Lagrangian-based count models that simultaneously accommodate flexible dispersion and excess zero observations.

Motivated by these considerations, this paper introduces the shifted Lagrangian exponential Poisson (SLEP) distribution together with its zero-inflated extension, namely the zero-inflated shifted Lagrangian exponential Poisson (ZISLEP) distribution. The proposed SLEP model is developed within the Lagrangian probability framework through a probability-generating mechanism that produces a positive-support count distribution with tractable probabilistic characteristics. The ZISLEP extension incorporates an additional structural-zero component to accommodate count data exhibiting excess zeros.

The proposed models possess tractable probabilistic properties, and explicit expressions for their PGFs, PMFs, moments, and dispersion measures are derived. These analytical results facilitate theoretical investigation, parameter estimation, and practical implementation.

This work fills an important gap in the count data literature. Existing Lagrangian probability distributions provide flexible probabilistic constructions but have received comparatively limited attention for developing models that simultaneously accommodate flexible dispersion and zero inflation within a unified framework. In contrast, generalized Poisson and related count models primarily achieve flexibility through modified mean–variance relationships, whereas conventional zero-inflated models rely on combining standard count distributions with a structural-zero component.

The proposed SLEP distribution addresses this gap by introducing a shifted Lagrangian generator in which the parameter c directly controls the probability of the smallest positive count while preserving the analytical tractability of the Lagrangian construction. The resulting family accommodates underdispersion, equidispersion, and overdispersion, while its zero-inflated extension, ZISLEP, simultaneously models excess zero observations within the same framework.

The contributions of this work are multifold. First, a new SLEP distribution and its zero-inflated extension (ZISLEP) are introduced within a unified Lagrangian framework. Second, explicit PGFs and PMFs are derived, together with analytical expressions for the moments and dispersion measures. Third, theoretical conditions governing underdispersion, equidispersion, and overdispersion

are established. Fourth, parameter estimation procedures based on both the method of moments and maximum likelihood estimation (MLE) are developed. Furthermore, regression extensions are proposed to accommodate covariate-dependent count data. Finally, the practical performance and flexibility of the proposed models are demonstrated through simulation studies and a real-data application.

The remainder of this paper is organized as follows. Section 2 presents the construction of the shifted Lagrangian distribution framework. Section 3 introduces the SLEP distribution and its zero-inflated extension. Section 4 studies important statistical properties such as the moments and dispersion properties. Sections 5 and 6 discuss the parameter estimation methods, such as MLE. The SLEP regression framework is developed in Section 7. The simulation study is given in Section 8. A real-data application is given in Section 9 to illustrate the practical applicability of the proposed model. Finally, concluding remarks are included in Section 10.

2. Shifted Lagrangian distribution construction

Traditional Lagrangian probability distributions are generated by PGFs satisfying the functional equation

$$\psi(u) = u\, g(\psi(u)), \tag{2.1}$$

where $\psi(u) = E(u^Y) = \sum_{y=0}^{\infty} P(Y = y)u^y$, with $|u| \leq 1$. Here, u denotes the argument of the PGF, while $g(\cdot)$ is the generating function appearing in the Lagrangian functional equation. Equation (2.1) is known as a Lagrangian functional equation because the unknown PGF appears on both sides through the composition $g(\psi(u))$. Solving this equation yields the PGF of the corresponding Lagrangian probability distribution.

The function $g(z)$ is assumed to be analytic in a neighborhood of the unit disk and to satisfy $g(1) = 1$; see [15, 16]. These conditions ensure that $\psi(u)$ is itself a valid PGF satisfying $\psi(1) = 1$. The condition $g(0) \neq 0$ guarantees a unique analytic solution $\psi(u)$ near $u = 0$ via Lagrange's inversion theorem.

To incorporate a structural baseline component, we introduce a shifted generator of the form

$$g(z) = c + (1 - c)h(z), \qquad 0 < c < 1, \tag{2.2}$$

where $h(z)$ is analytic near the unit disk and satisfies $h(1) = 1$ and $h(0) = 0$. Normalization is preserved since $g(1) = c + (1 - c)h(1) = 1$ and $g(0) = c \in (0, 1)$.

Let $z = \psi(u)$ denote the unique analytic solution of $z = u\, g(z)$. By the Lagrange inversion theorem (see [16]), the solution admits the expansion

$$\psi(u) = \sum_{y=1}^{\infty} \frac{u^y}{y!} \frac{d^{y-1}}{dz^{y-1}} [g(z)]^y \bigg|_{z=0}.$$

Hence, the PMF is given by

$$P(Y = y) = \frac{1}{y!} \frac{d^{y-1}}{dz^{y-1}} [g(z)]^y \bigg|_{z=0}, \qquad y = 1, 2, \ldots, \tag{2.3}$$

with $P(Y = 0) = 0$ following directly from $\psi(0) = 0$. On expanding $[g(z)]^y$ using the binomial theorem, it yields

$$[g(z)]^y = \sum_{k=0}^{y} \binom{y}{k} c^{y-k}(1-c)^k [h(z)]^k. \tag{2.4}$$

On substituting (2.4) into (2.3), it gives

$$P(Y = y) = \frac{1}{y!} \sum_{k=0}^{y} \binom{y}{k} c^{y-k}(1-c)^k \left. \frac{d^{y-1}}{dz^{y-1}} [h(z)]^k \right|_{z=0}.$$

For $y = 1$, we obtain $P(Y = 1) = c + (1-c)h(0) = c$, using $h(0) = 0$. For $y \geq 2$, note that $h(0) = 0$ implies that $[h(z)]^k$ has a zero of order at least k at $z = 0$. Hence, its $(y-1)$-th derivative at zero vanishes whenever $k \geq y$. The term $k = 0$ is constant and therefore also contributes zero after differentiation. Consequently, only indices $k = 1, \ldots, y-1$ contribute

$$P(Y = y) = \frac{1}{y!} \sum_{k=1}^{y-1} \binom{y}{k} c^{y-k}(1-c)^k \left. \frac{d^{y-1}}{dz^{y-1}} [h(z)]^k \right|_{z=0}, \qquad y = 2, 3, \ldots. \tag{2.5}$$

This representation separates the baseline parameter c from the mechanism governed by $h(z)$. The parameter c directly specifies the probability at $Y = 1$, since $P(Y = 1) = c$, while also influencing higher-order probabilities through the binomial weights c^{y-k}. The function $h(z)$ controls the higher-order probabilistic structure via the Lagrangian recursion.

In the next section, we specify a parametric form of $h(z)$ that generates a flexible family capable of capturing underdispersion, equidispersion, and overdispersion.

3. Shifted Lagrangian exponential Poisson family

In this section, we introduce a shifted Lagrangian generator with flexible dispersion properties, following the general construction in (2.2). The proposed construction uses an exponential modification of the shifted generator. The SLEP generator is defined as

$$g(z) = c + (1-c)ze^{\delta(z-1)}, \qquad 0 < c < 1,\ \delta > 0. \tag{3.1}$$

The normalization and nonvanishing conditions follow immediately from the general properties established in (2.2)

$$g(1) = c + (1-c) \cdot 1 \cdot e^0 = 1, \qquad g(0) = c + (1-c) \cdot 0 \cdot e^{-\delta} = c \in (0, 1).$$

Under the Lagrangian functional equation (2.1), setting $u = 0$ gives $\psi(0) = 0$, implying $P(Y = 0) = 0$ via (2.3). Hence, the model has no probability mass at zero. The parameter c determines the probability at $Y = 1$ via $P(Y = 1) = g(0) = c$, while also influencing higher-order probabilities through the Lagrangian structure, as seen in (2.5). Several limiting and special cases follow.

- If $\delta \to 0^+$, then $g(z) = c+(1-c)z$. The functional equation (2.1) becomes $\psi(u) = u(c+(1-c)\psi(u))$, yielding $\psi(u) = cu/[1-(1-c)u]$, which corresponds to a geometric distribution with $P(Y = y) = c(1-c)^{y-1},\ y = 1, 2, \ldots$

- If $c \to 0^+$, then $g(z) = ze^{\delta(z-1)}$. This corresponds to a Lagrangian exponential-Poisson construction with $P(Y = 1) = g(0) = 0$. Since $P(Y = 0) = 0$ and $P(Y = 1) = 0$, the support reduces to $\{2, 3, \ldots\}$.
- If $c \to 1^-$, then $g(z) \to 1$. The functional equation (2.1) reduces to $\psi(u) = u$, which is the PGF of a degenerate distribution concentrated at $Y = 1$.
- For $\delta > 0$, the exponential modification term $e^{\delta(z-1)}$ introduces additional dispersion flexibility relative to standard Poisson-type constructions.

Such properties make the SLEP family especially suitable for modeling count data with flexible dispersion and direct control of the probability at $Y = 1$ via the parameter c. The parameter δ gives additional control over the tail behavior and the concentration of probability mass around small counts, allowing the distribution to accommodate a wide range of dispersion patterns.

3.1. The SLEP PMF

The SLEP distribution is constructed within the Lagrangian framework using the functional equation in (2.1), where the generating function is defined in (3.1).

Theorem 1 (PMF of the SLEP distribution)**.** *Let Y be a discrete random variable with PGF given in* (2.1)*, where $g(z) = c + (1 - c)ze^{\delta(z-1)}$, with the parameters $0 < c < 1$ and $\delta > 0$. Then*

$$P_{\text{SLEP}}(Y = 0; c, \delta) = 0, \qquad P_{\text{SLEP}}(Y = 1; c, \delta) = c,$$

and for $y \geq 2$,

$$P_{\text{SLEP}}(Y = y; c, \delta) = \frac{1}{y} \sum_{k=1}^{y-1} \binom{y}{k} c^{y-k} (1 - c)^k e^{-\delta k} \frac{(\delta k)^{y-k-1}}{(y - k - 1)!}. \tag{3.2}$$

Proof. See Appendix A. □

To verify normalization, note that since $\psi(u)$ is a PGF, $\sum_{y\geq 1} P_{\text{SLEP}}(Y = y; c, \delta) = \psi(1)$. From the defining functional equation, $\psi(1) = g(\psi(1))$. Observe that $g(1) = c + (1 - c)e^0 = 1$, so 1 is a fixed point of g. Since a PGF satisfies $0 \leq \psi(1) \leq 1$, and the solution of the functional equation is the unique admissible solution in $[0, 1]$, it follows that $\psi(1) = 1$. Hence, $\sum_{y\geq 1} P_{\text{SLEP}}(Y = y; c, \delta) = 1$, which verifies that the PMF is properly normalized.

Since the SLEP distribution assigns zero probability to the origin, it is naturally suited for zero-truncated count data. To accommodate excess zeros, we introduce a zero-inflated extension.

3.2. Zero-inflated SLEP distribution

Although the SLEP distribution provides flexible dispersion modeling, its support excludes zero by construction, that is, $P_{\text{SLEP}}(Y = 0; c, \delta) = 0$.

However, many real count datasets exhibit excess zeros relative to standard count models. To accommodate such data, we define the ZISLEP distribution through a mixture representation as follows:

$$P_{\text{ZISLEP}}(Y = y; \pi, c, \delta) = \begin{cases} \pi + (1-\pi)\cdot 0 = \pi, & y = 0, \\ (1-\pi)P_{\text{SLEP}}(Y = y; c, \delta), & y = 1, 2, 3, \ldots, \end{cases} \tag{3.3}$$

where $0 \leq \pi \leq 1$ denotes the probability of structural zeros. The PGF of the ZISLEP distribution is $\psi_{\text{ZISLEP}}(u) = \pi + (1-\pi)\psi_{\text{SLEP}}(u)$.

Here, $\psi_{\text{SLEP}}(u)$ denotes the PGF of the SLEP distribution. The ZISLEP model in (3.3) follows the standard zero-inflated mixture framework, in which the additional parameter π modifies the distributional properties inherited from the SLEP component. Consequently, the moments and dispersion measures depend jointly on c, δ, and π.

In this formulation, the structural-zero component generates all zero observations, whereas the SLEP distribution models the conditional distribution of positive counts. Consequently, the support of the SLEP distribution on $\{1, 2, \ldots\}$ is fully consistent with its role as the positive count component of the ZISLEP model. In the real-data application, the SLEP model is therefore fitted only to the positive observations, while the ZISLEP model is fitted to the complete dataset.

The parameters of the proposed models have straightforward practical interpretations. The parameter c governs the probability of the smallest positive count through $P(Y = 1) = c$, so larger values of c increase the concentration of probability near one. The parameter δ influences the overall shape and dispersion of the distribution, affecting the spread of the probability mass and the heaviness of the upper tail. Consequently, different combinations of c and δ produce a wide range of distributional shapes. For the zero-inflated extension, the additional parameter π determines the proportion of structural zeros without altering the conditional distribution of the positive counts. Figure 1 illustrates these effects on the corresponding PMFs.

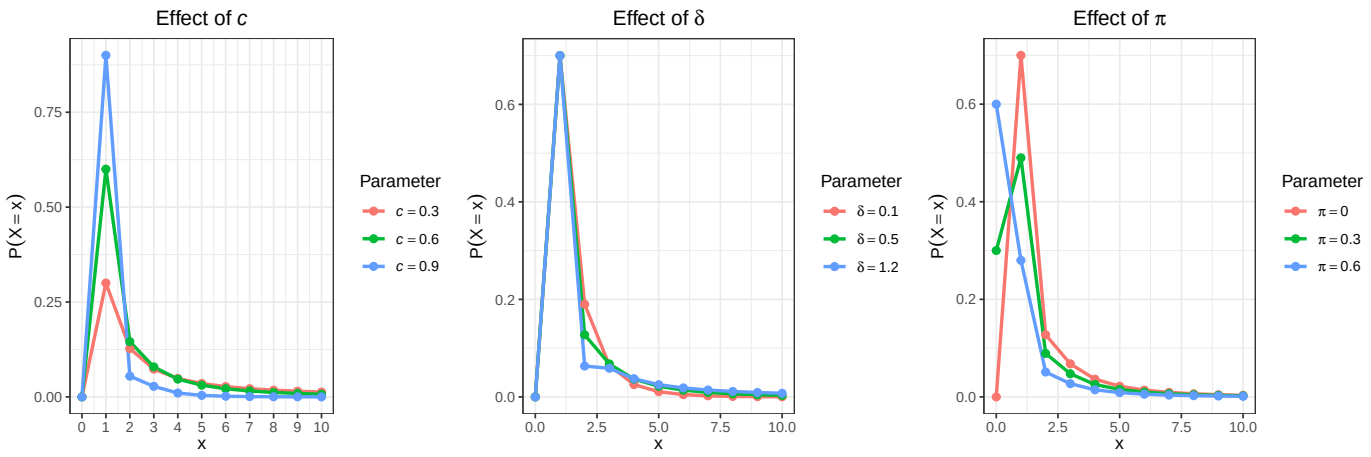

Figure 1. PMFs illustrating the effects of the SLEP and ZISLEP models' parameters. The three plots show the effect of varying c (with $\delta = 0.5$ fixed), varying δ (with $c = 0.7$ fixed), and varying the zero-inflation parameter π (with $c = 0.7$ and $\delta = 0.5$ fixed)

4. Moments and dispersion properties

Let $\psi(u)$ denote the PGF associated with the SLEP Lagrangian distribution defined in (3.1). The PGF satisfies the traditional Lagrangian functional equation (2.1). Under the standard Lagrangian

functional equation, $\psi(0) = 0$ implies $P(Y = 0) = 0$ via (2.3), so c is not the probability mass at zero but rather determines $P(Y = 1) = c$, consistent with the shifted generator construction in (2.2).

4.1. Fundamental properties

This subsection describes the main structural properties of the SLEP Lagrangian model, such as the normalization, the conditions of the existence of moments, and the properties of dispersion. These results give theoretical insight into the probabilistic behavior of the model and shed light on the role of the parameters (c, δ) in controlling the mean, variability, and tail behavior of the distribution.

Theorem 2 (Mean of the SLEP distribution)**.** *Let Y follow the SLEP Lagrangian distribution. A finite mean exists whenever the condition $(1 - c)(1 + \delta) < 1$ is met, and it is expressed as follows:*

$$\mathbb{E}[Y] = \mu = \frac{1}{1 - (1 - c)(1 + \delta)}. \tag{4.1}$$

Proof. See Appendix C. □

The mean in (4.1) depends jointly on the structural parameter c and the dispersion parameter δ. Increasing δ increases the expected count level and generally increases the variability as $(1-c)(1+\delta) \to 1^-$, leading to larger expected counts and heavier tails.

Theorem 3 (Variance of the SLEP distribution)**.** *If $(1 - c)(1 + \delta) < 1$, such that the first two moments exist, then the variance of Y is given by*

$$\mathrm{Var}(Y) = \frac{(1 - c)\left[c(1 + \delta)^2 + \delta\right]}{[1 - (1 - c)(1 + \delta)]^3}. \tag{4.2}$$

Proof. See Appendix D. □

The variance expression in (4.2) shows that both the parameter c and the exponential tilting parameter δ contribute directly to variability. Consequently, the SLEP model can capture heterogeneous dispersion structures in count data.

Theorem 4 (Index of dispersion)**.** *Provided that $(1 - c)(1 + \delta) < 1$, the index of dispersion is*

$$D = \frac{(1 - c)\left[c(1 + \delta)^2 + \delta\right]}{[1 - (1 - c)(1 + \delta)]^2}. \tag{4.3}$$

Proof. See Appendix E. □

The explicit form of the dispersion index in (4.3) demonstrates that the SLEP model possesses substantial flexibility in dispersion modeling through the joint effects of c and δ.

Proposition 1 (Dispersion regimes)**.** *Let D denote the dispersion index in Theorem 4, assuming $(1 - c)(1 + \delta) < 1$, such that the first two moments exist. Then*

- *$D > 1$ corresponds to overdispersion;*

- $D = 1$ *corresponds to equidispersion;*
- $D < 1$ *corresponds to underdispersion.*

The equidispersion curve in the (c, δ)*-parameter space is given by*

$$(1-c)\left[c(1+\delta)^2+\delta\right] = \left[1-(1-c)(1+\delta)\right]^2.$$

Proof. From the definition of the dispersion index and Theorem 4, we see that $D > 1$, $D = 1$, and $D < 1$ correspond to overdispersion, equidispersion, and underdispersion, respectively. Setting $D = 1$ in Theorem 4 yields the stated equidispersion curve. □

Therefore, the SLEP family provides a unified framework capable of modeling underdispersed, equidispersed, and overdispersed count data within a single parametric structure.

4.2. Moments of the ZISLEP distribution

Let $\mu_{\text{SLEP}} = \mathbb{E}_{\text{SLEP}}[Y]$ and $\sigma^2_{\text{SLEP}} = \text{Var}_{\text{SLEP}}(Y)$ denote the mean and variance of the SLEP distribution, respectively. Using the mixture representation in (3.3), the raw moments of the ZISLEP distribution are directly related to those of the SLEP's baseline distribution: $\mathbb{E}[Y] = (1-\pi)\mu_{\text{SLEP}}$ and $\mathbb{E}[Y^2] = (1-\pi)\left(\sigma^2_{\text{SLEP}} + \mu^2_{\text{SLEP}}\right)$. The variance is as follows:

$$\begin{aligned}\text{Var}(Y) &= \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\ &= (1-\pi)\left(\sigma^2_{\text{SLEP}} + \mu^2_{\text{SLEP}}\right) - (1-\pi)^2\mu^2_{\text{SLEP}} \\ &= (1-\pi)\sigma^2_{\text{SLEP}} + \pi(1-\pi)\mu^2_{\text{SLEP}}.\end{aligned}$$

If $D_{\text{SLEP}} = \sigma^2_{\text{SLEP}}/\mu_{\text{SLEP}}$ denotes the dispersion index of the SLEP distribution, then

$$\begin{aligned}D_{\text{ZISLEP}} &= \frac{\text{Var}(Y)}{\mathbb{E}[Y]} \\ &= \frac{(1-\pi)\sigma^2_{\text{SLEP}} + \pi(1-\pi)\mu^2_{\text{SLEP}}}{(1-\pi)\mu_{\text{SLEP}}} \\ &= \frac{\sigma^2_{\text{SLEP}}}{\mu_{\text{SLEP}}} + \pi\mu_{\text{SLEP}} \\ &= D_{\text{SLEP}} + \pi\mu_{\text{SLEP}}.\end{aligned}$$

Hence, the dispersion index of the ZISLEP distribution is obtained by an additive adjustment to the baseline dispersion index, $D_{\text{ZISLEP}} = D_{\text{SLEP}} + \pi\mu_{\text{SLEP}}$. Thus, the zero-inflation parameter π contributes directly to additional variability, increasing dispersion relative to the baseline SLEP model.

5. Parameter estimation

Parameter estimation for the SLEP Lagrangian distribution can be performed using both the method of moments and maximum likelihood. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the distribution.

5.1. Method of moments estimation

Define the sample mean and sample variance as

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2.$$

Equating the sample moments with the theoretical moments from Theorems 2 and 3 yields

$$\bar{Y} = \mu, \qquad S^2 = \mathrm{Var}(Y), \tag{5.1}$$

where the explicit expressions for μ and $\mathrm{Var}(Y)$ are given in (4.1) and (4.2), respectively. The moment equations are valid under the existence condition $(1-c)(1+\delta) < 1$, which guarantees finite moments.

5.1.1. Estimation procedures

From the mean equation $\bar{Y} = \mu$, solving for c using (4.1) yields

$$\hat{c} = 1 - \frac{1 - \frac{1}{\bar{Y}}}{1+\delta} = 1 - \frac{\bar{Y} - 1}{(1+\delta)\bar{Y}}.$$

Since the support of the SLEP distribution is $\{1, 2, \ldots\}$, we have $\bar{Y} \geq 1$. For nondegenerate samples with $\bar{Y} > 1$, the estimator satisfies $\hat{c} \in (0, 1)$.

From the same mean equation, solving for δ when c is known yields

$$\hat{\delta} = \frac{1 - \frac{1}{\bar{Y}}}{1-c} - 1.$$

To ensure $\hat{\delta} > 0$, the sample mean must satisfy $\bar{Y} > 1/c$. When both c and δ are unknown, the system in (5.1) is solved jointly.

Closed-form expressions are available when one parameter is assumed to be known. Otherwise, numerical methods are required to solve the nonlinear moment equations jointly. These estimates can also serve as the initial values for MLE.

6. Likelihood function and MLE

An alternative approach for estimating the parameters c and δ of the SLEP distribution is the method of MLE.

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the SLEP distribution. Let $y_1, \ldots, y_n$ denote the observed sample values and

$$I = \{\, i \in \{1, \ldots, n\} : y_i \geq 2 \,\}.$$

The likelihood function for (c, δ) is given by

$$L(c, \delta) = c^{n_1} \prod_{i \in I} \left(\frac{1}{y_i} \sum_{k=1}^{y_i - 1} \binom{y_i}{k} c^{y_i - k} (1-c)^k e^{-\delta k} \frac{(\delta k)^{y_i - k - 1}}{(y_i - k - 1)!} \right), \tag{6.1}$$

where

$$n_1 = \sum_{i=1}^{n} \mathbb{I}(y_i = 1)$$

is the number of observations equal to 1. The log-likelihood function corresponding to (6.1) is given by

$$\ell(c, \delta) = \log L(c, \delta) = n_1 \log c + \sum_{i \in I} \left[-\log y_i + \log \left(S_i(c, \delta) \right) \right], \tag{6.2}$$

where

$$S_i(c, \delta) = \sum_{k=1}^{y_i - 1} \binom{y_i}{k} c^{y_i - k} (1 - c)^k e^{-\delta k} \frac{(\delta k)^{y_i - k - 1}}{(y_i - k - 1)!}.$$

The score functions are given by

$$\frac{\partial \ell}{\partial c} = \frac{n_1}{c} + \sum_{i \in I} \frac{1}{S_i(c, \delta)} \cdot \frac{\partial S_i(c, \delta)}{\partial c}, \qquad \frac{\partial \ell}{\partial \delta} = \sum_{i \in I} \frac{1}{S_i(c, \delta)} \cdot \frac{\partial S_i(c, \delta)}{\partial \delta}, \tag{6.3}$$

where, using the SLEP PMF in (3.2), we have

$$\frac{\partial S_i(c, \delta)}{\partial c} = \sum_{k=1}^{y_i - 1} \binom{y_i}{k} e^{-\delta k} \frac{(\delta k)^{y_i - k - 1}}{(y_i - k - 1)!} \cdot \left[(y_i - k) c^{y_i - k - 1} (1 - c)^k - k c^{y_i - k} (1 - c)^{k - 1} \right],$$

$$\frac{\partial S_i(c, \delta)}{\partial \delta} = \sum_{k=1}^{y_i - 1} \binom{y_i}{k} c^{y_i - k} (1 - c)^k \cdot e^{-\delta k} (\delta k)^{y_i - k - 1} \left[\frac{y_i - k - 1}{\delta} - k \right] \frac{1}{(y_i - k - 1)!}.$$

Setting $\partial \ell / \partial c = 0$ and $\partial \ell / \partial \delta = 0$ produces a nonlinear system of likelihood equations that does not admit closed-form solutions because of the summation terms contained in $S_i(c, \delta)$. Consequently, the parameter estimates are obtained numerically using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton algorithm. Initial values are obtained from the method of moments estimators described in Section 5.1. Throughout the simulation study and the real-data application, the optimization algorithm converged successfully for all fitted datasets. A step-by-step MLE algorithm is provided in Appendix G, Algorithm 2.

Standard errors for the MLE are obtained from the observed Fisher information matrix; details are provided in Appendix F.

7. SLEP regression model

The SLEP distribution can be extended to a regression framework for modeling count responses influenced by the explanatory variables. Let $Y_i \sim \text{SLEP}(c_i, \delta_i)$, $i = 1, 2, \ldots, n$, where the model parameters depend on explanatory variables through suitable link functions.

Under the SLEP model, from Theorem 1, we have $P_{\text{SLEP}}(Y_i = 0; c_i, \delta_i) = 0$ and $P_{\text{SLEP}}(Y_i = 1; c_i, \delta_i) = c_i$, where c_i specifies the probability mass at $Y_i = 1$, while δ_i governs the dispersion behavior

through its effect on the variance structure established in Theorem 3. Let $\mathbf{u}_i = (1, u_{i1}, u_{i2}, \ldots, u_{ip})^\top$ and $\mathbf{v}_i = (1, v_{i1}, v_{i2}, \ldots, v_{iq})^\top$ denote covariate vectors associated with the i-th observation, where $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ are the vectors of regression coefficients corresponding to $\mathbf{u}_i$ and $\mathbf{v}_i$, respectively.

The parameter c_i is modeled using a logistic link function

$$\log\left(\frac{c_i}{1-c_i}\right) = \mathbf{u}_i^\top \boldsymbol{\beta} \quad \Longrightarrow \quad c_i = \frac{1}{1+\exp(-\mathbf{u}_i^\top \boldsymbol{\beta})}. \tag{7.1}$$

This specification guarantees $0 < c_i < 1$.

To accommodate heterogeneity in the count responses, the dispersion parameter is modeled as

$$\log(\delta_i) = \mathbf{v}_i^\top \boldsymbol{\tau} \quad \Longrightarrow \quad \delta_i = \exp(\mathbf{v}_i^\top \boldsymbol{\tau}). \tag{7.2}$$

This specification guarantees $\delta_i > 0$.

7.1. Likelihood function for SLEP regression

Let $\boldsymbol{\Theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\tau}^\top)^\top$ denote the vector of regression parameters. Assuming independence, the likelihood function is

$$L(\boldsymbol{\Theta}) = \prod_{i=1}^{n} P_{\text{SLEP}}\,(y_i \mid \mathbf{u}_i, \mathbf{v}_i; \boldsymbol{\Theta}).$$

The corresponding log-likelihood function is

$$\ell(\boldsymbol{\Theta}) = \sum_{i=1}^{n} \log P_{\text{SLEP}}\,(y_i \mid \mathbf{u}_i, \mathbf{v}_i; \boldsymbol{\Theta}).$$

Since $P_{\text{SLEP}}(Y_i = 0; c_i, \delta_i) = 0$, the SLEP regression model is appropriate for positive count data in which zero responses are not observed. Applications involving zero counts would require an extension of the model, such as the ZISLEP formulation, to accommodate zero-valued responses.

For $y_i \geq 1$, the PMF is obtained from Theorem 1 by replacing c and δ with their covariate-dependent forms c_i and δ_i. Because the resulting likelihood function is nonlinear in the regression parameters, estimation requires numerical optimization procedures such as Newton–Raphson, Fisher scoring, or quasi-Newton methods.

The score vector and observed Fisher information matrix are obtained by combining the derivatives of the link functions with the likelihood structure of the SLEP model; detailed expressions are provided in Appendix F.1.

7.2. ZISLEP regression model with covariates

To increase the flexibility and practical applicability of the proposed framework, we extend the ZISLEP distribution to a regression setting where all three parameters are allowed to depend on covariates.

Let Y_i denote the response variable for observation i and let $\mathbf{z}_i$, $\mathbf{x}_i$, and $\mathbf{w}_i$ be covariate vectors associated with the zero-inflation, shape, and dispersion components, respectively. The ZISLEP regression model is defined as

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \mathbf{z}_i^\top \boldsymbol{\gamma}, \quad \log\left(\frac{c_i}{1-c_i}\right) = \mathbf{x}_i^\top \boldsymbol{\beta}, \quad \log(\delta_i) = \mathbf{w}_i^\top \boldsymbol{\eta}.$$

Equivalently

$$\pi_i = (1 + \exp(-\mathbf{z}_i^\top \boldsymbol{\gamma}))^{-1}, \quad c_i = (1 + \exp(-\mathbf{x}_i^\top \boldsymbol{\beta}))^{-1}, \quad \delta_i = \exp(\mathbf{w}_i^\top \boldsymbol{\eta}).$$

It is important to note that c_i is a parameter of the positive-count SLEP component and should not be interpreted as the marginal probability that $Y_i = 1$. Since the ZISLEP model is a mixture, the unconditional probability of observing one count is

$$P_{\text{ZISLEP}}(Y_i = 1; \pi_i, c_i, \delta_i) = (1 - \pi_i)c_i,$$

whereas $P_{\text{SLEP}}(Y_i = 1; c_i, \delta_i) = c_i$ within the positive-count SLEP component.

Under this formulation, the PMF of Y_i is given by

$$P_{\text{ZISLEP}}(Y_i = y_i \mid \mathbf{z}_i, \mathbf{x}_i, \mathbf{w}_i) = \begin{cases} \pi_i, & y_i = 0, \\ (1 - \pi_i)\, P_{\text{SLEP}}(Y_i = y_i; c_i, \delta_i), & y_i \geq 1. \end{cases} \tag{7.3}$$

The covariate vectors $\mathbf{z}_i$, $\mathbf{x}_i$, and $\mathbf{w}_i$ may be distinct or may share common explanatory variables, depending on the application. The proposed formulation does not require the same covariates to be used for all three model components, thereby providing considerable modeling flexibility. In practice, parsimonious model specifications are generally preferred to reduce the model's complexity and facilitate estimation.

The log-likelihood function for a sample $(y_1, \ldots, y_n)$ is given by

$$\begin{aligned} \ell(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\eta}) = &\sum_{i=1}^{n} \mathbb{I}(y_i = 0) \log(\pi_i) \\ &+ \sum_{i=1}^{n} \mathbb{I}(y_i \geq 1) \left[\log(1 - \pi_i) + \log P_{\text{SLEP}}(Y_i = y_i; c_i, \delta_i)\right]. \end{aligned}$$

7.3. MLE for ZISLEP

Let $y_1, \ldots, y_n$ be a random sample from the ZISLEP regression model and $I = \{i \in \{1, 2, \ldots, n\} : y_i \geq 2\}$. Define

$$n_0 = \sum_{i=1}^{n} \mathbb{I}(y_i = 0), \qquad n_1 = n - n_0.$$

For the special case without covariates (i.e., $\pi_i = \pi$, $c_i = c$, $\delta_i = \delta$), the log-likelihood simplifies to

$$\ell(\pi, c, \delta) = n_0 \log(\pi) + n_1 \log(1 - \pi) + \sum_{i \in I} \log P_{\text{ZISLEP}}(Y = y_i; \pi, c, \delta), \tag{7.4}$$

where $P_{\text{ZISLEP}}(Y = y_i; \pi, c, \delta)$ denotes the ZISLEP PMF defined in (3.3). The likelihood separates into a Bernoulli component for π and a conditional SLEP component for (c, δ) based on the positive observations.

The score equation for π is given by

$$\frac{\partial \ell}{\partial \pi} = \frac{n_0}{\pi} - \frac{n_1}{1 - \pi}. \tag{7.5}$$

Setting the score equation (7.5) to zero and solving for π, it follows that

$$\hat{\pi} = \frac{n_0}{n_0 + n_1}. \tag{7.6}$$

Since $n = n_0 + n_1$ (total number of observations, where n_0 and n_1 are the counts for $y_i = 0$ and $y_i \geq 1$, respectively), (7.6) yields the closed-form estimator $\hat{\pi} = n_0/n$. Moreover, because $n_0 \sim \text{binomial}(n, \pi)$, the estimator is unbiased and consistent.

Substituting $\hat{\pi}$ into the simplified log-likelihood (7.4) gives the maximized log-likelihood

$$\ell_p(c, \delta) = n_0 \log\left(\frac{n_0}{n}\right) + n_1 \log\left(1 - \frac{n_0}{n}\right) + \sum_{i \in I} \log P_{\text{SLEP}}(Y = y_i; c, \delta). \tag{7.7}$$

Since the first two terms in (7.7) do not depend on (c, δ), maximizing $\ell_p(c, \delta)$ is equivalent to maximizing

$$\ell_p^*(c, \delta) = \sum_{i \in I} \log P_{\text{SLEP}}(Y = y_i; c, \delta), \tag{7.8}$$

The conditional distribution of the positive observations corresponds to a zero-truncated SLEP distribution, since conditioning on $Y_i \in \{1, 2, 3, \dots\}$ removes the structural zero component.

Maximization of (7.8) requires numerical optimization due to the complexity of the SLEP PMF in (3.2).

For the special case without covariates, the estimator of π is available in closed form as given in (7.6), whereas (c, δ) are obtained by numerically maximizing the profile log-likelihood in (7.8). For the general regression model, all regression coefficients are estimated simultaneously by maximizing the full log-likelihood using the procedure described in Section 6.

Standard errors are computed from the observed Fisher information matrix as described in Appendix F.

7.4. Model selection and diagnostics

Model adequacy for the SLEP regression model may be assessed using likelihood-based information criteria together with residual diagnostics. These tools provide complementary approaches for evaluating model fit, complexity, and predictive performance.

7.4.1. Information criteria

For model comparison, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) are defined as

$$\text{AIC} = -2\ell(\hat{\boldsymbol{\Theta}}) + 2k, \qquad \text{BIC} = -2\ell(\hat{\boldsymbol{\Theta}}) + k\log(n), \tag{7.9}$$

where $\ell(\hat{\boldsymbol{\Theta}})$ in (7.9) denotes denotes the log-likelihood evaluated at the maximum likelihood estimator $\hat{\boldsymbol{\Theta}}$, k is the number of estimated parameters, and n is the sample size.

Smaller values of AIC and BIC indicate a better balance between goodness-of-fit and model complexity.

Residual diagnostics may also be performed using randomized quantile residuals for discrete-response models. Such residuals are approximately standard normal under correct model specification and may be useful for identifying departures from the model's assumptions.

7.4.2. Goodness-of-fit assessment

A goodness-of-fit assessment may be performed using the Pearson chi-square statistic based on grouped observed and expected frequencies as follows:

$$\chi^2 = \sum_{j=1}^{m} \frac{(O_j - E_j)^2}{E_j}, \tag{7.10}$$

where O_j and E_j denote the observed and fitted frequencies for the j-th count category, respectively, and m denotes the total number of categories used in the goodness-of-fit test. The corresponding degrees of freedom are $df = m - 1 - k$, where k denotes the number of estimated parameters.

To satisfy the assumptions of the Pearson goodness-of-fit test, categories with small expected frequencies are combined before computing the statistic to ensure that all expected cell frequencies are at least 5.

Under the null hypothesis that the fitted model adequately describes the data, the statistic χ^2 in (7.10) is approximately chi-square distributed with df degrees of freedom. Larger p-values indicate no evidence against the fitted model, whereas smaller p-values suggest a lack of fit.

8. Simulation study

A simulation study was conducted to explore the finite sample behavior of maximum likelihood estimators for the SLEP distribution with various parameter settings and sample sizes. The accuracy of the estimators, their variability, interval estimation, and convergence behavior are studied.

8.1. Simulation design

Random samples from the SLEP distribution were generated using inverse transform sampling based on the cumulative distribution function obtained from the SLEP PMF. Maximum likelihood estimates were computed using the BFGS quasi-Newton optimization algorithm with suitable starting values. Details of the data-generation and estimation algorithms are provided in Appendix G.

Four simulation scenarios were considered to represent different dispersion regimes: Near equidispersion, moderate overdispersion, underdispersion, and heavy overdispersion. The parameter configurations are summarized in Table 1.

Table 1. Simulation parameter configurations and dispersion regimes.

Scenario	c	δ	D	Dispersion type
I	0.81	0.60	1.05	Near equidispersion
II	0.60	0.25	1.90	Moderate overdispersion
III	0.85	0.10	0.24	Underdispersion (elevated mass at $Y = 1$)
IV	0.50	0.30	4.67	Heavy overdispersion

The dispersion index D was computed using the expression derived in Theorem 4. Figure 2 displays the corresponding PMFs under the four scenarios.

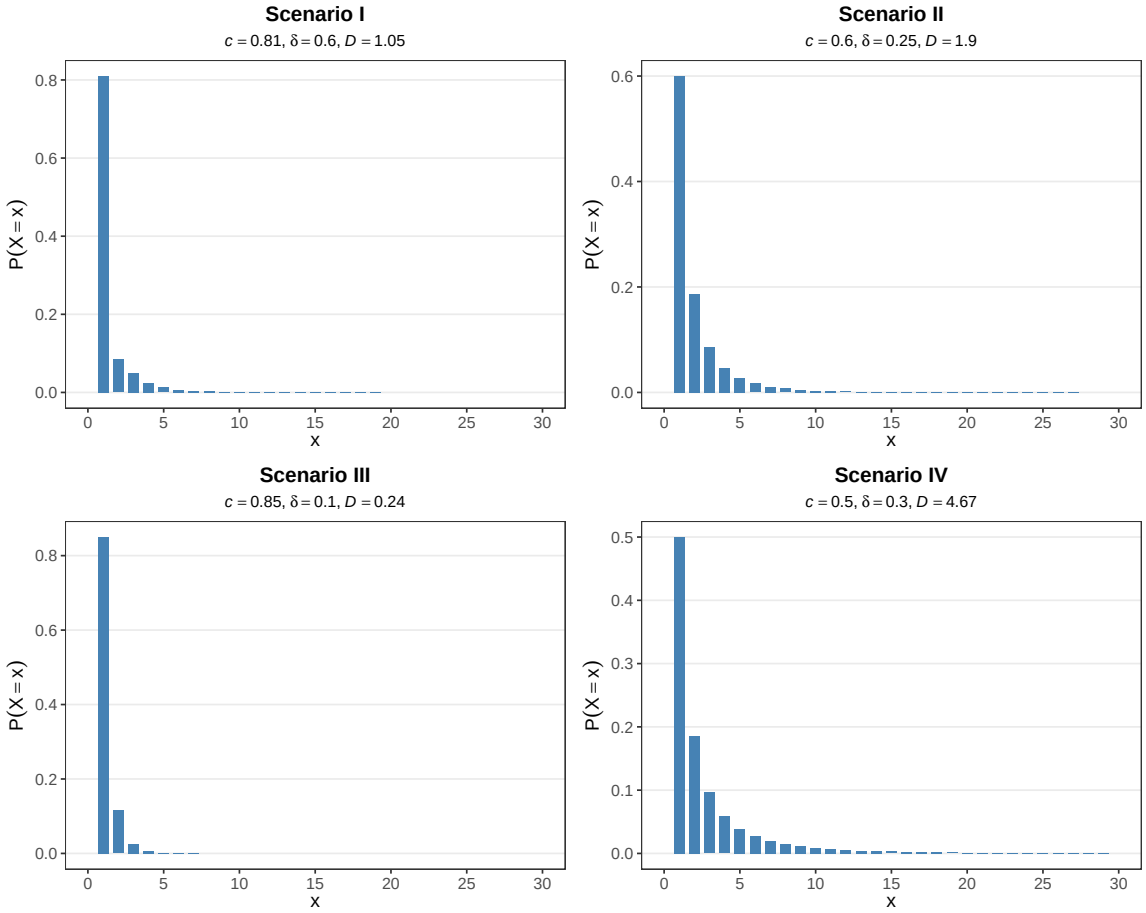

Figure 2. PMFs of the SLEP distribution under four dispersion scenarios. The figure compares the shapes of the PMFs corresponding to near equidispersion (Scenario I), moderate overdispersion (Scenario II), underdispersion (Scenario III), and heavy overdispersion (Scenario IV), illustrating the flexibility of the proposed distribution across different dispersion regimes.

8.2. Performance measures

Let $\boldsymbol{\theta}_0 = (c, \delta)^\top$ denote the true parameter vector. For each simulation setting, we generate R independent replications, yielding the maximum likelihood estimators

$$\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_R, \qquad \hat{\boldsymbol{\theta}}_r = (\hat{c}_r, \hat{\delta}_r)^\top,$$

which are obtained via the numerical optimization procedure described in Algorithm 2.

For a generic component $\theta \in \{c, \delta\}$ with a true value θ_0, the empirical bias is defined as

$$\text{Bias}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}(\hat{\theta}_r - \theta_0),$$

and the relative bias is given by

$$\text{RelBias}(\hat{\theta}) = \frac{\text{Bias}(\hat{\theta})}{\theta_0} \times 100.$$

The mean squared error (MSE) and root mean squared error (RMSE) are defined as

$$\mathrm{MSE}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}(\hat{\theta}_r - \theta_0)^2, \qquad \mathrm{RMSE}(\hat{\theta}) = \sqrt{\mathrm{MSE}(\hat{\theta})}.$$

The empirical SD is given by

$$\mathrm{SD}(\hat{\theta}) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}(\hat{\theta}_r - \bar{\hat{\theta}})^2}, \qquad \bar{\hat{\theta}} = \frac{1}{R}\sum_{r=1}^{R}\hat{\theta}_r.$$

In addition, inference is based on the observed Fisher information matrix defined in Appendix F (see (F.1)). For each replication r, let $\mathcal{I}_O(\hat{\boldsymbol{\theta}}_r)$ denote the observed Fisher information matrix evaluated at $\hat{\boldsymbol{\theta}}_r$. The corresponding estimated covariance matrix is

$$\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}}_r) = \mathcal{I}_O(\hat{\boldsymbol{\theta}}_r)^{-1}.$$

The asymptotic standard error of a parameter $\theta \in \{c, \delta\}$ in replication r is obtained from the diagonal elements of this matrix as follows:

$$\widehat{\mathrm{SE}}_r(\hat{c}) = \sqrt{\left[\mathcal{I}_O(\hat{\boldsymbol{\theta}}_r)^{-1}\right]_{11}}, \qquad \widehat{\mathrm{SE}}_r(\hat{\delta}) = \sqrt{\left[\mathcal{I}_O(\hat{\boldsymbol{\theta}}_r)^{-1}\right]_{22}}.$$

The average asymptotic standard error over all replications is then computed as

$$\overline{\mathrm{SE}}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}\widehat{\mathrm{SE}}_r(\hat{\theta}).$$

For interval estimation, Wald-type confidence intervals with the nominal level $1 - \alpha = 0.95$ are constructed as

$$\hat{\theta}_r \pm z_{1-\alpha/2}\,\widehat{\mathrm{SE}}_r(\hat{\theta}),$$

where $z_{0.975} = 1.96$. The empirical coverage probability is defined as

$$\mathrm{Coverage} = \frac{1}{R}\sum_{r=1}^{R}\mathbb{I}(\theta_0 \in \mathrm{CI}_r),$$

where $\mathbb{I}(\cdot)$ denotes the indicator function and

$$\mathrm{CI}_r = \left[\hat{\theta}_r - z_{1-\alpha/2}\widehat{\mathrm{SE}}_r(\hat{\theta}),\ \hat{\theta}_r + z_{1-\alpha/2}\widehat{\mathrm{SE}}_r(\hat{\theta})\right].$$

An effective estimation procedure is expected to yield small bias, relative bias, MSE, and RMSE, together with coverage probabilities close to the nominal level, indicating reliable finite-sample performance under the SLEP model.

8.3. Results and discussion

Simulation experiments were conducted for sample sizes $n \in \{100, 250, 500, 1000\}$ with $R = 1000$ replications per scenario. Maximum likelihood estimates were obtained using the BFGS quasi-Newton algorithm described in Algorithm 2. Estimation performance was assessed using the empirical bias, relative bias (%), RMSE, empirical SD, average asymptotic standard error (AvgSE), empirical coverage probability (Cov.Prob) for nominal 95% Wald-type intervals, and convergence rate (Conv.Rate). Across all settings, the optimization routine converged successfully in every replication, yielding Conv.Rate = 1.000.

Table 2 to Table 5 summarize the results for the four scenarios. In general, empirical bias and RMSE decrease monotonically with increasing n, providing evidence consistent with the asymptotic consistency of the MLE.

8.3.1. Scenario I: Near equidispersion

Estimators exhibit negligible bias across all sample sizes, with absolute relative bias below 0.09% for c and below 1.4% for δ. RMSE decreases from 0.039 to 0.013 for c and from 0.222 to 0.073 for δ as n increases from 100 to 1000. AvgSE values closely match the empirical SDs, and the coverage probabilities range from 0.926 to 0.946 for c and 0.928 to 0.940 for δ, remaining close to the nominal 95% level. Overall, the confidence intervals are slightly liberal at smaller sample sizes but approach the nominal level as n increases.

Table 2. Simulation results for Scenario I (near equidispersion).

n	Parameter	Bias	RelBias (%)	RMSE	SD	AvgSE	Cov.Prob
100	c	-0.0007	-0.086	0.0388	0.0388	0.0381	0.926
	δ	-0.0083	-1.391	0.2223	0.2224	0.2227	0.928
250	c	0.0006	0.075	0.0250	0.0250	0.0240	0.930
	δ	0.0051	0.854	0.1469	0.1470	0.1417	0.936
500	c	0.0002	0.021	0.0170	0.0170	0.0170	0.946
	δ	0.0049	0.816	0.1026	0.1026	0.0997	0.940
1000	c	0.0001	0.017	0.0127	0.0127	0.0120	0.928
	δ	0.0069	1.157	0.0729	0.0726	0.0705	0.940

8.3.2. Scenario II: moderate overdispersion

Absolute relative bias remains below 1.7% across all sample sizes. RMSE decreases systematically, with δ improving from 0.127 at $n = 100$ to 0.038 at $n = 1000$. Coverage probabilities range from 0.910 to 0.970; for δ, coverage is slightly below nominal at $n = 100$ (0.910) but improves with sample size. Thus, the confidence intervals are mildly liberal for small samples and become approximately nominal for moderate and large samples.

Table 3. Simulation results for Scenario II (Moderate overdispersion).

n	Parameter	Bias	RelBias (%)	RMSE	SD	AvgSE	Cov.Prob
100	c	0.0005	0.083	0.0488	0.0488	0.0487	0.940
	δ	0.0042	1.678	0.1271	0.1271	0.1231	0.910
250	c	0.0007	0.116	0.0309	0.0309	0.0309	0.946
	δ	-0.0006	-0.230	0.0792	0.0793	0.0771	0.940
500	c	0.0001	0.012	0.0201	0.0202	0.0218	0.970
	δ	0.0014	0.557	0.0539	0.0539	0.0543	0.948
1000	c	-0.0017	-0.284	0.0145	0.0144	0.0155	0.966
	δ	-0.0030	-1.206	0.0376	0.0376	0.0382	0.942

8.3.3. Scenario III: Underdispersion with elevated mass at $Y = 1$

This setting is more challenging for the estimation of δ. For c, the absolute relative bias remains below 0.2% across all sample sizes. For δ, the relative bias is 15.7% at $n = 100$ but decreases to 2.0% at $n = 1000$. Coverage for δ improves from 0.714 at $n = 100$ to 0.950 at $n = 1000$, while coverage for c remains close to nominal across all sample sizes. Hence, the confidence intervals for δ are markedly liberal in small samples but reach the nominal level as the sample size increases.

Table 4. Simulation results for Scenario III (underdispersion with elevated mass at $Y = 1$).

n	Parameter	Bias	RelBias (%)	RMSE	SD	AvgSE	Cov.Prob
100	c	0.0008	0.092	0.0318	0.0318	0.0345	0.944
	δ	0.0157	15.743	0.1287	0.1278	0.1108	0.714
250	c	-0.0014	-0.161	0.0216	0.0216	0.0224	0.964
	δ	0.0087	8.693	0.0820	0.0816	0.0768	0.863
500	c	-0.0007	-0.078	0.0151	0.0151	0.0159	0.960
	δ	0.0005	0.524	0.0585	0.0585	0.0570	0.946
1000	c	-0.0001	-0.016	0.0114	0.0114	0.0113	0.944
	δ	0.0020	2.049	0.0417	0.0416	0.0411	0.950

8.3.4. Scenario IV: heavy overdispersion

Despite increased variability, the estimators remain stable. Absolute relative bias is below 0.1% for c and below 1% for δ. RMSE for δ decreases from 0.124 at $n = 100$ to 0.035 at $n = 1000$. Coverage for δ improves from 0.896 at $n = 100$ to 0.948 at $n = 1000$, while coverage for c ranges from 0.904 to 0.954. Therefore, the confidence intervals are slightly liberal for small sample sizes and become approximately nominal for larger samples.

Table 5. Simulation results for Scenario IV (Heavy overdispersion).

n	Parameter	Bias	RelBias (%)	RMSE	SD	AvgSE	Cov.Prob
100	c	0.0001	0.014	0.0534	0.0535	0.0492	0.904
	δ	-0.0029	-0.977	0.1242	0.1242	0.1134	0.896
250	c	-0.0000	-0.006	0.0305	0.0305	0.0313	0.952
	δ	-0.0005	-0.181	0.0753	0.0754	0.0718	0.934
500	c	0.0005	0.092	0.0219	0.0219	0.0222	0.954
	δ	0.0002	0.071	0.0497	0.0497	0.0509	0.960
1000	c	0.0000	0.005	0.0152	0.0152	0.0157	0.954
	δ	-0.0002	-0.064	0.0350	0.0350	0.0360	0.948

8.3.5. Discussion of the simulation results

The simulation results demonstrate the satisfactory finite-sample performance of the proposed estimation procedure across all dispersion regimes. Estimators for c exhibit negligible bias and stable coverage properties throughout all scenarios. Estimation of δ becomes more challenging under underdispersion and small sample sizes, where larger relative bias and slightly lower coverage probabilities are observed. This behavior is expected because the parameter δ governs the dispersion and subtle changes in the upper tail of the distribution. When the sample size is small, relatively few moderate and large counts are observed, making δ more difficult to estimate accurately than c, which is primarily determined by the frequency of observations equal to one. However, these effects diminish as the sample size increases, with both the relative bias and RMSE decreasing substantially.

The empirical standard deviations become increasingly consistent with the corresponding average estimated standard errors as the sample size increases, indicating the improved finite-sample accuracy of the maximum likelihood estimators. Moreover, the empirical coverage probabilities remain close to the nominal 95% confidence level for most scenarios, suggesting that confidence intervals based on the observed Fisher information provide satisfactory finite-sample performance. The close agreement between the empirical standard deviations and the average estimated standard errors further supports the adequacy of the observed Fisher information for statistical inference. MLE was used because of its desirable asymptotic properties, including consistency, asymptotic normality, and efficiency under standard regularity conditions. In addition, it facilitates likelihood-based model comparison through the information criteria defined in (7.9).

Across all scenarios, RMSE decreases with increasing sample size, and the optimization algorithm converges successfully in all replications. Figure 3 visually confirms this decreasing trend for both parameters across all four scenarios. Overall, the results suggest that the proposed SLEP estimation framework provides reliable and computationally stable inference for a wide range of count data settings.

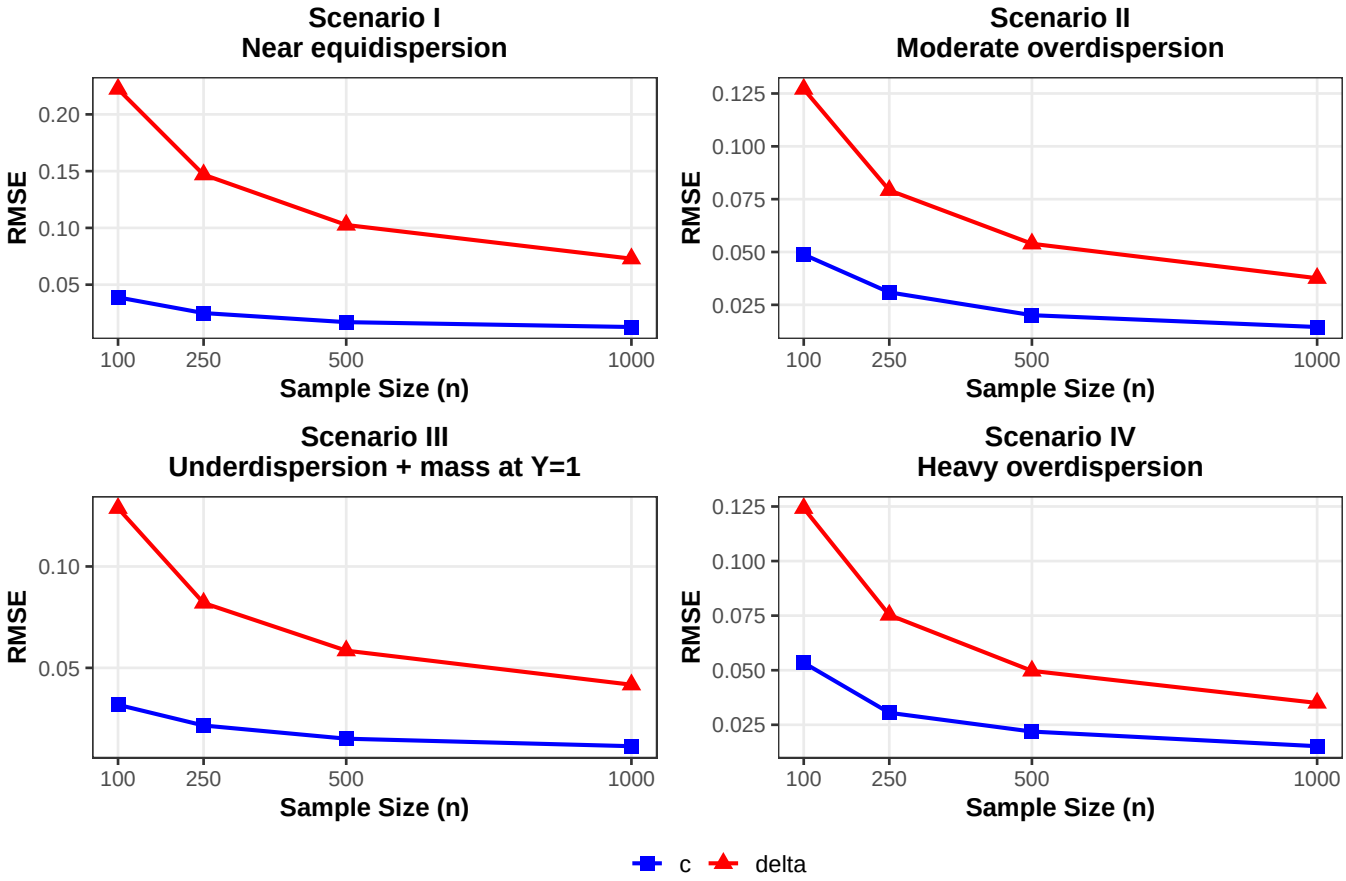

Figure 3. RMSE comparison across the four simulation scenarios. The plots compare the RMSEs of the maximum likelihood estimators for the parameters c (blue squares) and δ (red triangles) as functions of the sample size n. In all scenarios, the RMSE decreases monotonically as n increases, providing empirical evidence supporting the asymptotic consistency of the maximum likelihood estimators.

9. Application

We illustrate the proposed ZISLEP model using a real count dataset characterized by excess zeros and overdispersion. The analysis compares ZISLEP with standard competing models, namely Poisson, NB, and ZIP, using likelihood-based information criteria and goodness-of-fit measures. All parameter estimates are obtained by maximum likelihood using the BFGS optimization procedure.

Since the SLEP distribution assigns zero probability at the origin ($P(X = 0) = 0$), it cannot be fitted to the complete dataset. Therefore, SLEP is fitted only to the positive observations ($n = 1049$) and is compared with other zero-truncated count models, namely the truncated Poisson and truncated NB distributions. Consequently, the full-data models (Poisson, NB, ZIP, and ZISLEP) and the zero-truncated models are evaluated separately throughout this section.

9.1. Data description

The empirical analysis uses the DoctorVisits dataset from the Applied Econometrics with R (AER) package, where the response variable represents the number of doctor visits. The dataset contains 5190 observations.

The sample mean and variance are 0.302 and 0.637, respectively, yielding a variance-to-mean ratio of 2.111 and suggesting overdispersion relative to the Poisson model. In addition, approximately 79.79% of observations are zeros, indicating substantial zero inflation. The observed counts range from 0 to 9.

9.2. Model comparison

Table 6 reports the log-likelihood, AIC, and BIC values for the competing models. The Poisson model provides the poorest fit, indicating that the data substantially depart from the equidispersion assumption. Incorporating overdispersion through the NB model considerably improves the model's fit, while the ZIP model accounts for excess zeros but remains less competitive.

Table 6. Model comparison for the DoctorVisits dataset.

Model	Log-likelihood	AIC	BIC
Full-data models (n = 5190)			
Poisson	−3983.194	7968.389	7974.943
NB	−3585.992	7175.983	7189.092
ZIP	−3714.236	7432.471	7445.580
ZISLEP (Proposed)	**−3557.971**	**7121.943**	**7141.606**
Zero-truncated models (n = 1049)			
Truncated Poisson	−1101.970	2205.941	2210.896
Truncated NB	−958.214	1920.428	1930.339
SLEP (Proposed)	**−945.706**	**1895.412**	**1905.323**

Note: Bold values indicate the best model within each group of competing models.

Table 6 presents the likelihood-based comparison for both groups of competing models. Among the full-data models, the proposed ZISLEP model achieves the largest log-likelihood and the smallest AIC and BIC values, indicating the best overall fit to the DoctorVisits data. For the positive observations, the proposed SLEP model also provides the largest log-likelihood and the smallest AIC and BIC values, outperforming both the truncated Poisson and truncated NB models. These results indicate that the proposed models provide greater flexibility in modeling both zero inflation and the distribution of positive counts.

9.3. Goodness-of-fit analysis

Table 7 reports Pearson chi-square goodness-of-fit statistics for the fitted models. Although all models are formally rejected at conventional significance levels because of the large sample size, the statistics remain useful for relative comparison.

Table 7. Chi-square goodness-of-fit comparison for competing models.

Model	χ^2	df	p-value
Full-data models ($n = 5190$)			
Poisson	2845.535	3	$< 1e-08$
NB	122.063	5	$< 1e-08$
ZIP	454.262	3	$< 1e-08$
ZISLEP (proposed)	**43.939**	5	$\mathbf{2.4e-08}$
Zero-truncated models ($n = 1049$)			
Truncated Poisson	454.262	3	$< 1e-08$
Truncated NB	88.101	4	$< 1e-08$
SLEP (proposed)	**43.858**	5	$\mathbf{2.5e-08}$

Note: Bold values indicate the model with the smallest chi-square statistic within each group.

The zero-truncated models were fitted only to the positive observations and are therefore compared only with other zero-truncated models, whereas the full-data models were fitted to the complete dataset, including zero counts. Consequently, the goodness-of-fit results should be interpreted separately for the two groups of models.

For the zero-truncated models, the proposed SLEP distribution yields the smallest chi-square statistic, indicating the closest agreement between the observed and fitted frequencies. The truncated Poisson model produces the same chi-square statistic as the ZIP model because, conditional on positive counts, the ZIP distribution reduces exactly to a zero-truncated Poisson distribution with the same Poisson mean. Consequently, both models induce identical fitted probabilities for the positive observations.

Among the full-data models, the Poisson model provides the poorest fit, while the NB and ZIP models substantially reduce the discrepancy between the observed and fitted frequencies. The proposed ZISLEP model yields the smallest chi-square statistic within this group, indicating the best overall fit to the complete dataset.

The ZISLEP model combines a zero-inflation component with a SLEP distribution for positive counts. The parameter π controls the probability of structural zeros, c regulates the probability mass at $Y = 1$, and δ governs the dispersion behavior of positive counts. This structure enables simultaneous modeling of both excess zeros and overdispersion within a unified framework.

9.4. Bootstrap analysis

To assess the stability of the proposed estimation procedure, a bootstrap analysis based on 1000 resamples was conducted for both the proposed ZISLEP model and the corresponding SLEP model fitted to the positive observations (see Table 8).

Table 8. Bootstrap performance of the MLEs for the DoctorVisits data.

Model	Parameter	Estimate	Bias	Rel. Bias (%)	RMSE	SD	Avg. SE	CP	95% CI
ZISLEP	π	0.798	−0.00003	−0.004	0.006	0.006	0.006	0.930	[0.786, 0.809]
	c	0.740	0.00094	0.128	0.013	0.013	0.013	0.940	[0.717, 0.771]
	δ	0.269	0.00068	0.252	0.040	0.040	0.041	0.955	[0.197, 0.343]
SLEP	c	0.740	0.00090	0.001	0.013	0.013	0.013	0.940	[0.717, 0.771]
	δ	0.269	0.00098	0.004	0.040	0.040	0.041	0.955	[0.197, 0.343]

The bootstrap results indicate stable estimation performance. Bias values remain negligible for all parameters, while empirical standard deviations closely match the average estimated standard errors. Coverage probabilities are reasonably close to the nominal 95% level, suggesting adequate finite-sample performance of the inferential procedure.

Overall, the empirical analysis demonstrates that the proposed ZISLEP model provides the best fit among the full-data models, while the proposed SLEP model provides the best fit among the zero-truncated models. These results indicate that incorporating the flexible SLEP distribution into both the zero-inflated and zero-truncated modeling frameworks substantially improves the modeling of count data exhibiting excess zeros and overdispersion.

10. Concluding remarks

In this paper, we introduce the SLEP distribution, a flexible discrete model within the framework of Lagrangian generating functions. The SLEP distribution is defined on the positive integers, naturally excluding zero, and allows for various dispersion behaviors through its parameters: c controls the probability mass at $X = 1$, and δ manages tail behavior and dispersion. Mathematically rigorous justification is provided by the derivation using Lagrange inversion in order to ensure that the PMF is well-defined and correctly normalized.

The SLEP framework extends naturally to more complex settings, including the ZISLEP model for count data with excess zeros and regression formulations that incorporate covariates through logistic and log-link functions. Simulation studies indicate that the maximum likelihood estimators perform well for finite samples under a range of dispersion scenarios, including near equidispersion, moderate and heavy overdispersion, and underdispersion. Estimation of δ is more difficult under underdispersion with small sample sizes, but these effects are alleviated by larger ones. An application on the DoctorVisits dataset confirms the practical utility of the ZISLEP model. It outperforms Poisson, NB, and ZIP alternatives in terms of the AIC, BIC, and chi-square goodness-of-fit.

There are several areas of future research. First, the SLEP framework could be extended to the multivariate case to model dependence among correlated count variables, e.g., in epidemiology and insurance. Second, extensions of regression with different link functions and regularized estimation (e.g., the least absolute shrinkage and selection operator (LASSO) and ridge regression) would allow for the variable selection in high-dimensional settings. Finally, additional theoretical results such as bias correction and refined asymptotic approximations would improve the statistical properties and practical reliability of the model.

The proposed SLEP and ZISLEP models extend the family of Lagrangian count distributions by introducing a shifted generating mechanism that provides direct control over the probability of

the smallest positive count while retaining analytical tractability. This offers practitioners a flexible alternative to existing generalized Poisson and zero-inflated count models for analyzing data exhibiting varying dispersion levels and excess zeros.

Appendices

A. Proof of Theorem 1

Proof. For a PGF satisfying (2.1), the Lagrange inversion formula gives (2.3). Consider $g(z)$ in (3.1), where $h(z) = ze^{\delta(z-1)}$. By raising $g(z)$ to the power y and applying the binomial expansion, we obtain

$$[g(z)]^y = \sum_{k=0}^{y} \binom{y}{k} c^{y-k}(1-c)^k [h(z)]^k.$$

For $k \geq 1$, note that $[h(z)]^k = z^k e^{\delta k(z-1)} = e^{-\delta k} z^k e^{\delta kz}$. Using the power-series expansion of the exponential function

$$e^{\delta kz} = \sum_{j=0}^{\infty} \frac{(\delta k)^j}{j!} z^j,$$

we obtain

$$[h(z)]^k = e^{-\delta k} \sum_{j=0}^{\infty} \frac{(\delta k)^j}{j!} z^{k+j}.$$

The coefficient of z^{y-1} occurs when $k + j = y - 1$, that is, $j = y - k - 1$. Hence, we have

$$D^{y-1}[h(z)]^k\Big|_{z=0} = (y-1)!e^{-\delta k} \frac{(\delta k)^{y-k-1}}{(y-k-1)!},$$

for $k = 1, \ldots, y-1$. Substituting into the inversion formula (2.3) yields

$$P(Y = y) = \frac{(y-1)!}{y!} \sum_{k=1}^{y-1} \binom{y}{k} c^{y-k}(1-c)^k e^{-\delta k} \frac{(\delta k)^{y-k-1}}{(y-k-1)!}.$$

Since $(y-1)!/y! = 1/y$, we obtain

$$P(Y = y) = \frac{1}{y} \sum_{k=1}^{y-1} \binom{y}{k} c^{y-k}(1-c)^k e^{-\delta k} \frac{(\delta k)^{y-k-1}}{(y-k-1)!}, \qquad y = 2, 3, \ldots$$

For $y = 1$, $P(Y = 1) = g(0) = c$. Hence, the required PMF follows. This completes the proof of Theorem 1.

□

B. Derivatives of the generator

Lemma 1 (First derivative of the generator)**.** *For the SLEP generator $g(z)$ in* (3.1)*, its first derivative with respect to z at $z = 1$ is $g'(1) = (1 - c)(1 + \delta)$, where the prime denotes differentiation with respect to z.*

Proof. Let $h(z) = ze^{\delta(z-1)}$. Then its first derivative with respect to z is

$$\begin{aligned} h'(z) &= e^{\delta(z-1)} + \delta z e^{\delta(z-1)} \\ &= e^{\delta(z-1)}(1 + \delta z). \end{aligned}$$

Evaluating at $z = 1$, we obtain $h'(1) = 1 + \delta$. Since $g(z) = c + (1 - c)h(z)$, differentiating it gives $g'(z) = (1 - c)h'(z)$. Therefore, we have

$$g'(1) = (1 - c)h'(1) = (1 - c)(1 + \delta).$$

This completes the proof. □

Lemma 2 (Second derivative of the generator)**.** *For the SLEP generator $g(z)$ in* (3.1)*, its second derivative with respect to z at $z = 1$ is $g''(1) = (1 - c)\delta(\delta + 2)$, where the double prime denotes differentiation with respect to z.*

Proof. Recall that $h(z) = ze^{\delta(z-1)}$. From Lemma 1, we have $h'(z) = e^{\delta(z-1)}(1 + \delta z)$. Differentiating $h'(z)$ once more, we obtain

$$\begin{aligned} h''(z) &= \delta e^{\delta(z-1)}(1 + \delta z) + \delta e^{\delta(z-1)} \\ &= \delta e^{\delta(z-1)}(\delta z + 2). \end{aligned}$$

Evaluating at $z = 1$, we obtain $h''(1) = \delta(\delta + 2)$. Since $g(z) = c + (1 - c)h(z)$, differentiating twice gives $g''(z) = (1 - c)h''(z)$. Therefore, we have

$$g''(1) = (1 - c)\delta(\delta + 2).$$

This completes the proof.

□

C. Proof of Theorem 2

Proof. For Lagrangian distributions, a generic formula for the mean is given by

$$\mu = \frac{1}{1 - g'(1)},$$

where $g'(1) < 1$ (see [16]). By Lemma 1, we have $g'(1) = (1 - c)(1 + \delta)$. Therefore, we have

$$\mu = \frac{1}{1 - (1 - c)(1 + \delta)},$$

subject to the existence condition $(1 - c)(1 + \delta) < 1$. This completes the proof.

□

D. Proof of Theorem 3

Proof. For Lagrangian distributions, a generic formula for the variance is given by

$$\mathrm{Var}(Y) = \frac{g''(1) + g'(1) - [g'(1)]^2}{[1 - g'(1)]^3},$$

where $g'(1) < 1$ (see [16]). From Lemmas 1 and 2, we have

$$g'(1) = (1 - c)(1 + \delta), \qquad g''(1) = (1 - c)\delta(\delta + 2).$$

Let $a = g'(1) = (1 - c)(1 + \delta)$ and $b = g''(1) = (1 - c)\delta(\delta + 2)$. Then, we have

$$\mathrm{Var}(Y) = \frac{b + a - a^2}{(1 - a)^3}, \tag{D.1}$$

which exists provided $a < 1$, i.e., $(1 - c)(1 + \delta) < 1$.

Substituting and simplifying the numerator in (D.1), we have

$$\begin{aligned} b + a - a^2 &= (1 - c)\delta(\delta + 2) + (1 - c)(1 + \delta) - (1 - c)^2(1 + \delta)^2 \\ &= (1 - c)\left[\delta(\delta + 2) + (1 + \delta) - (1 - c)(1 + \delta)^2\right]. \end{aligned}$$

Since $\delta(\delta + 2) + (1 + \delta) = (1 + \delta)^2 + \delta$, we obtain

$$b + a - a^2 = (1 - c)\left[(1 + \delta)^2 + \delta - (1 - c)(1 + \delta)^2\right].$$

Hence, we have

$$b + a - a^2 = (1 - c)\left[c(1 + \delta)^2 + \delta\right].$$

Therefore

$$\mathrm{Var}(Y) = \frac{(1 - c)\left[c(1 + \delta)^2 + \delta\right]}{[1 - (1 - c)(1 + \delta)]^3},$$

subject to the existence condition $(1 - c)(1 + \delta) < 1$. This completes the proof.

□

E. Proof of Theorem 4

Proof. From Theorems 2 and 3, the index of dispersion is defined as

$$D = \frac{\mathrm{Var}(Y)}{\mu}.$$

Since

$$\mu = \frac{1}{1 - (1 - c)(1 + \delta)},$$

and

$$\mathrm{Var}(Y) = \frac{(1-c)\left[c(1+\delta)^2+\delta\right]}{[1-(1-c)(1+\delta)]^3},$$

we obtain

$$D = \frac{(1-c)\left[c(1+\delta)^2+\delta\right]}{[1-(1-c)(1+\delta)]^2}.$$

□

F. Fisher information and asymptotic inference

Let $\boldsymbol{\theta} = (c, \delta)^\top$ denote the parameter vector, and let $ell(\boldsymbol{\theta})$ be the log-likelihood function defined in (6.2) for a random sample $X_1, \ldots, X_n$. The Fisher information matrix is defined as

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}_{\boldsymbol{\theta}}\left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right].$$

Because the SLEP likelihood contains summation terms embedded within logarithmic functions, closed-form expressions for $\mathcal{I}(\boldsymbol{\theta})$ are not available. Consequently, inference is based on the observed Fisher information matrix

$$\mathcal{I}_O(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}. \tag{F.1}$$

The asymptotic covariance matrix of the maximum likelihood estimator $\hat{\boldsymbol{\theta}} = (\hat{c}, \hat{\delta})^\top$ is approximated by

$$\widehat{\mathrm{Var}}(\hat{\boldsymbol{\theta}}) = \mathcal{I}_O(\hat{\boldsymbol{\theta}})^{-1}.$$

Hence, the approximate standard errors are obtained from the diagonal elements as follows:

$$\widehat{\mathrm{SE}}(\hat{c}) = \sqrt{\left[\mathcal{I}_O(\hat{\boldsymbol{\theta}})^{-1}\right]_{11}}, \qquad \widehat{\mathrm{SE}}(\hat{\delta}) = \sqrt{\left[\mathcal{I}_O(\hat{\boldsymbol{\theta}})^{-1}\right]_{22}}.$$

Under standard regularity conditions, we have

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta})^{-1}\right),$$

which establishes the asymptotic normality of the maximum likelihood estimator and provides the theoretical basis for large-sample statistical inference.

F.1. Score functions for SLEP regression

Let $\mathbf{u}_i$ and $\mathbf{v}_i$ denote the covariate vectors for the c_i and δ_i components, respectively. From the link functions in (7.1) and (7.2), we have

$$\frac{\partial c_i}{\partial \boldsymbol{\beta}} = c_i(1-c_i)\mathbf{u}_i, \qquad \frac{\partial \delta_i}{\partial \boldsymbol{\tau}} = \delta_i \mathbf{v}_i.$$

Let $\ell_i = \log P(Y_i = y_i \mid \mathbf{u}_i, \mathbf{v}_i)$ denote the individual log-likelihood contribution. Then the score functions are given by

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \frac{\partial \ell_i}{\partial c_i} c_i(1 - c_i)\mathbf{u}_i, \qquad \frac{\partial \ell}{\partial \boldsymbol{\tau}} = \sum_{i=1}^{n} \frac{\partial \ell_i}{\partial \delta_i} \delta_i \mathbf{v}_i.$$

The partial derivatives $\partial \ell_i / \partial c_i$ and $\partial \ell_i / \partial \delta_i$ follow directly from the SLEP PMF in (3.2). Due to the summation form of the PMF, closed-form expressions are generally unavailable and these derivatives are evaluated numerically within the optimization routine.

The full score vector is therefore

$$\mathbf{S}(\boldsymbol{\Theta}) = \left(\frac{\partial \ell}{\partial \boldsymbol{\beta}}, \frac{\partial \ell}{\partial \boldsymbol{\tau}} \right)^{\top},$$

and the likelihood equations $\mathbf{S}(\boldsymbol{\Theta}) = \mathbf{0}$ are solved numerically.

G. Simulation algorithms

G.1. Data generation procedure

Random samples from the SLEP distribution are generated using inverse transform sampling based on the PMF in Theorem 1. The CDF is computed numerically as

$$F_y = \sum_{j=0}^{y} P(Y = j),$$

and observations are generated via inverse transform sampling.

Algorithm 1 Generation of SLEP random variates

Require: Parameters c, δ, sample size N
Ensure: Sample $Y_1, \ldots, Y_N$
1: Compute $P(Y = y)$ for $y = 0, \ldots, y_{\max}$ with $P(Y = 0) = 0$
2: Choose $y_{\max}$ such that $\sum_{y=1}^{y_{\max}} P(Y = y) \approx 1$
3: Construct $F_y = \sum_{j=0}^{y} P(Y = j)$
4: **for** $i = 1$ to N **do**
5: Draw $U \sim U(0, 1)$
6: Set $Y_i = \min\{y : F_y \geq U\}$
7: **end for**
8: **return** $\{Y_i\}_{i=1}^{N}$

G.2. MLE procedure

Let $\boldsymbol{\theta} = (c, \delta)^{\top}$ and $\ell(\boldsymbol{\theta})$ be the log-likelihood in (6.2). Estimation is performed using BFGS optimization. The initial values are based on sample moments

$$\hat{c}^{(0)} = \frac{n_1}{n}, \qquad \hat{\delta}^{(0)} = \frac{1 - \frac{1}{\bar{y}}}{1 - \hat{c}^{(0)}} - 1,$$

with $n_1 = \sum \mathbb{I}(y_i = 1)$ and $\bar{y} = n^{-1} \sum y_i$. The algorithm iterates until convergence of the log-likelihood or gradient norm below a tolerance ε.

Algorithm 2 Maximum likelihood estimation for SLEP

Require: Data $y_1, \ldots, y_n$, tolerance ε, max iterations T
Ensure: MLE $\hat{\boldsymbol{\theta}}$
Compute n_1 and $\bar{y}$
Initialize $\boldsymbol{\theta}^{(0)} = (\hat{c}^{(0)}, \hat{\delta}^{(0)})$
for $k = 1$ to T **do**
    Evaluate $\ell(\boldsymbol{\theta}^{(k-1)})$
    Update $\boldsymbol{\theta}^{(k)}$ via BFGS
    Compute $g_k = \|\nabla \ell(\boldsymbol{\theta}^{(k)})\|$
    if $g_k < \varepsilon$ **then**
        break
    end if
end for
return $\boldsymbol{\theta}^{(k)}$

Use of Generative-AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgment

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Small Research Groups Program under grant number RGP1/38/47. We appreciate the helpful comments, which enhanced the clarity, accuracy, and quality of this manuscript.

Conflict of interest

The author declares that there is no conflict of interest.

References

1. R. Winkelmann, *Econometric analysis of count data*, 5 Eds., Springer, Berlin/Heidelberg, Germany, 2008. http://dx.doi.org/10.1007/978-3-540-78389-3
2. A. C. Cameron, P. K. Trivedi, *Regression analysis of count data*, 2 Eds., Cambridge University Press, Cambridge, UK, 2013. http://dx.doi.org/10.1017/CBO9780511814365
3. J. M. Hilbe, *Modeling count data*, Cambridge University Press, Cambridge, UK, 2014. http://dx.doi.org/10.1017/CBO9781139236065

4. S. D. Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, Bachelier, Paris, France, 1837.

5. Z. Yang, J. W. Hardin, C. L. Addy, Q. H. Vuong, Testing approaches for overdispersion in Poisson regression versus the generalized Poisson model, *Biometrical J.,* **49** (2007), 565–584. http://dx.doi.org/10.1002/bimj.200610340

6. B. A. Desmarais, J. J. Harden, Testing for zero inflation in count models: Bias correction for the Vuong test, *Stata J.,* **13** (2013), 810–835. http://dx.doi.org/10.1177/1536867X1301300408

7. M. Ridout, J. Hinde, C. G. B. Demétrio, A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives, *Biometrics,* **57** (2001), 219–223. http://dx.doi.org/10.1111/j.0006-341X.2001.00219.x

8. J. M. Hilbe, *Negative binomial regression*, 2 Eds., Cambridge University Press, Cambridge, UK, 2011. http://dx.doi.org/10.1017/CBO9780511973420

9. D. Lambert, Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, **34** (1992), 1–14. http://dx.doi.org/10.2307/1269547

10. W. H. Greene, *Accounting for excess zeros and sample selection in Poisson and negative binomial regression models*, Department of Economics Working Paper EC-94-10, Stern School of Business, New York University, 1994. Available from: `https://archive.nyu.edu/handle/2451/26263`.

11. P. C. Consul, G. C. Jain, A generalization of the Poisson distribution, *Technometrics,* **15** (1973), 791–799. http://dx.doi.org/10.1080/00401706.1973.10489112

12. F. Famoye, K. P. Singh, Zero-inflated generalized Poisson regression model with an application to domestic violence data, *J. Data Sci.,* **4** (2006), 117–130. http://dx.doi.org/10.6339/JDS.2006.04(1).257

13. R. Berk, J. MacDonald, Overdispersion and Poisson regression, *J. Quant. Criminol.,* **24** (2008), 269–284. http://dx.doi.org/10.1007/s10940-008-9048-4

14. A. Zeileis, C. Kleiber, S. Jackman, Regression models for count data in R, *J. Stat. Softw.,* **27** (2008), 1–25. http://dx.doi.org/10.18637/jss.v027.i08

15. P. C. Consul, L. R. Shenton, Use of Lagrange expansion for generating discrete generalized probability distributions, *SIAM J. Appl. Math.,* **23** (1972), 239–248. http://dx.doi.org/10.1137/0123026

16. P. C. Consul, F. Famoye, *Lagrangian probability distributions*, Birkhäuser, Boston, MA, USA, 2006.

17. G. Dattoli, S. Lorenzutta, D. Sacchetti, Two-variable Lagrange expansion and new families of mixed generating functions, *Ann. Univ. Ferrara*, **45** (1999), 87–90. http://dx.doi.org/10.1007/BF02825946

18. N. L. Johnson, S. Kotz, A. W. Kemp, *Univariate discrete distributions*, 3 Eds., Wiley, New York, NY, USA, 2005. http://dx.doi.org/10.1002/0471715816

19. F. A. A. Aldhufairi, Rational-power shifted Lagrangian distribution for count data with flexible dispersion, *Mathematics,* **14** (2026), 1673. http://dx.doi.org/10.3390/math14101673

20. F. A. A. Aldhufairi, H. A. A. Alolaywi, The modified Sunil distribution: Theory, estimation, and comparative applications using batterers and victim survey data, *J. Stat. Theory Pract.,* **19** (2025), 23. http://dx.doi.org/10.1007/s42519-025-00425-7

AIMS Press