The new Poisson mixed weighted Lindley distribution with applications to insurance claims data

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Songklanakarin Journal of Science and Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>SJST-2018-0224.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>06-Oct-2018</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Atikankul, Yupapin; Kasetsart University Faculty of Science, Statistics; Rajamangala University of Technology Phra Nakhon, Mathematics and Statistics; Thongteeraparp, Ampai; Kasetsart University, Statistics; Bodhisuwan, Winai; Kasetsart University, Department of Statistics</td>
</tr>
<tr>
<td>Keyword:</td>
<td>count data, overdispersion, new weighted Lindley distribution, mixed Poisson distribution, maximum likelihood estimation</td>
</tr>
</tbody>
</table>
Original Article

The new Poisson mixed weighted Lindley distribution

with applications to insurance claims data

Yupapin Atikankul¹, Ampai Thongteeraparp¹,
and Winai Bodhisuwan¹*

¹Department of Statistics, Faculty of Science, Kasetsart University,
10900, Thailand

* Corresponding author, Email address: fsciwnb@ku.ac.th

Abstract

Mixed Poisson distributions have been applied for overdispersed count data analysis. In this paper, an alternative mixed Poisson distribution is proposed. The proposed distribution is derived by mixing the Poisson distribution with the new weighted Lindley distribution, named as the new Poisson mixed weighted Lindley distribution. Some special cases and several statistical properties have been derived including shape, factorial moments, probability generating function, moment generating function and moments. The maximum likelihood estimators of the parameters are obtained. Finally, two automobile insurance claims data sets are analyzed to compare the performance of the proposed distribution with some competitive distributions.

Keywords: count data, overdispersion, new weighted Lindley distribution, mixed Poisson distribution, maximum likelihood estimation
1. Introduction

Count data models are utilized in various fields such as public health, insurance, agriculture, etc. Some applications of count data have been studied, for example, the daily numbers of seizures of the patients with epilepsy (Albert, 1991), the number of automobile insurance claims in Germany (Tröbliger, 1961) and the number of roots of type of apple tree rootstock (Ridout, Demétrio, & Hinde, 1998). The Poisson distribution is a classic distribution for describing count data with the property of equality of the variance and mean, i.e. equidispersion. Unfortunately, in practical count data, the variance is usually larger than the mean, referred to as overdispersion, and rarely, the variance may be smaller than the mean, underdispersion. The Poisson distribution cannot be applied to account for these phenomena.

Overdispersion occurs in almost all count data. An approach that has been widely applied to deal with overdispersed count data is that of mixed Poisson distributions (Karlis & Xekalaki, 2005; Grandell, 1997; Gupta & Ong, 2005). The best known distribution of the Poisson type is the negative binomial distribution (Greenwood & Yule, 1920), which it arises from a mixture of the Poisson and gamma distributions.

The Lindley distribution (Lindley, 1958) arises from mixing the exponential distribution with the gamma distribution. Ghitany, Atieh, and Nadarajah (2008) applied the Lindley distribution to a data set of waiting times for bank customers. The result showed that the Lindley distribution provided a better fit than the exponential distribution. Thus in various papers, the Lindley distribution and some of its modifications have been studied and developed as mixing distributions for a mixed Poisson distribution. Sankaran (1970) proposed the discrete Poisson-Lindley distribution, which arises from the Poisson distribution and the Lindley mixing
distribution (Lindley, 1958). Mahmoudi and Zakerzadeh (2010) introduced a mixture of
the Poisson and generalized Lindley distributions (Zakerzadeh & Dolati, 2009).
Shanker, Sharma, and Shanker (2012) proposed a mixed Poisson distribution with the
two parameter Lindley mixing distribution (Shanker, Sharma, & Shanker, 2013).
Shanker and Mishra (2014) introduced the two-parameter Lindley distribution (Shanker
& Mishra, 2013) as a mixing distribution. In the same year, the Poisson–weighted
Lindley distribution was proposed by El-Monsef and Sohsah (2014); Manesh, Hamzah,
and Zamani (2014). Wongrin and Bodhisuwan (2016) proposed a mixture of the
Poisson and new generalized Lindley distributions (Elbatal, Merovci, & Elgarhy, 2013).

The new weighted Lindley (NWL) distribution was proposed by Asgharzadeh,
Bakouch, Nadarajah, and Sharafi (2016). The NWL distribution is obtained by mixing
the weighted exponential distribution (Gupta & Kundu, 2009) with the weighted gamma
distribution. The probability density function (pdf) is

\[
g(x) = \frac{\theta^2(1+\alpha)^2}{\alpha \theta(1+\alpha) + \alpha(2+\alpha)(1-x)(1-e^{-\theta x})} e^{-\theta x},
\]  

for \( x > 0, \theta > 0 \) and \( \alpha > 0 \).

The Lindley and weighted Lindley distributions are special cases of the NWL
distribution. The pdf is log-concave and unimodal. A data set of the amount of carbon in
leaves from the different mountainous areas of Navarra, Spain was modeled with the
NWL distribution. The result showed that the NWL distribution is a better fit than the
compared distributions.

In this paper, the new Poisson mixed weighted Lindley (NPWL) distribution, a
mixture of the Poisson and the new weighted Lindley distributions, is proposed. The
rest of this paper is arranged as follows. In Section 2, the NPWL distribution is
introduced. Some important statistical properties such as shape, factorial moments,
probability generating function, moment generating function and moments are exhibited in Section 3. In Section 4, the steps for random variate generation are presented. In Section 5, the maximum likelihood estimators of the parameters are discussed. In Section 6, the NPWL distribution is applied to some automobile insurance claims data sets. The conclusions are presented in Section 7.

2. The new Poisson mixed weighted Lindley distribution

The proposed distribution is a mixture of the Poisson distribution and the new weighted Lindley distribution (Asgharzadeh, Bakouch, Nadarajah, & Sharafi, 2016). Moreover; some special cases, the distribution function and the survival function of the proposed distribution are shown in this section.

Let $X | \lambda$ be a Poisson random variable with probability mass function (pmf)

$$f(x | \lambda) = \frac{e^{-\lambda} \lambda^x}{x!},$$

for $x = 0, 1, 2, \ldots$ and $\lambda > 0$, written as $X | \lambda \sim \text{Pois}(\lambda)$. Now we assume that $\lambda$ is distributed as the new weighted Lindley distribution with pdf

$$g(\lambda) = \frac{\theta^2(1+\alpha)^2}{a\theta(1+\alpha)+\alpha(2+\alpha)}(1+\lambda)(1-e^{-\theta\lambda})e^{-\theta\lambda},$$

for $\lambda > 0, \theta > 0$ and $\alpha > 0$, written as $\lambda \sim \text{NWL}(\theta, \alpha)$; then the unconditional discrete random variable $X$ follows the new Poisson mixed weighted Lindley distribution.

Proposition 1 A discrete random variable $X$ is distributed as the NPWL distribution, with the shape and scale parameters are $\theta$ and $\alpha$, respectively, denoted as $X \sim \text{NPWL}(\theta, \alpha)$, the pmf of $X$ is
\[ f(x) = \frac{\theta^2(1+\alpha)^2}{\alpha \theta(1+\alpha) + \alpha(2+\alpha)} \left( \frac{\theta + x + 2}{(\theta + 1)^{x+2}} - \frac{\alpha \theta + x + 2}{(\alpha \theta + 1)^{x+2}} \right), \] (3)

for \( x = 0, 1, 2, \ldots \), where \( \theta > 0 \) and \( \alpha > 0 \).

**Proof**

Let \( X|\lambda \sim \text{Pois}(\lambda) \) and \( \lambda \sim \text{NWL}(\theta, \alpha) \), the pmf of \( X \) is derived as the following

\[
    f(x) = \int_0^\infty f(x|\lambda)g(\lambda)d\lambda
\]

\[
    = \int_0^\infty \frac{\theta^2(1+\alpha)^2}{\alpha \theta(1+\alpha) + \alpha(2+\alpha)} \left( 1 - e^{-\theta \alpha \lambda} \right) e^{-\theta \lambda} \frac{\lambda^x e^{-\lambda}}{x!} d\lambda
\]

\[
    = \frac{\theta^2(1+\alpha)^2}{(\alpha \theta(1+\alpha) + \alpha(2+\alpha))x!} \left( \int_0^\infty \lambda^{x+1} e^{-(\theta + 1)\lambda} d\lambda + \int_0^\infty \lambda^x e^{-(\theta + 1)\lambda} d\lambda - \int_0^\infty \lambda^{x+1} e^{-(\alpha \theta + \theta + 1)\lambda} d\lambda - \int_0^\infty \lambda^x e^{-(\alpha \theta + \theta + 1)\lambda} d\lambda \right)
\]

\[
    = \frac{\theta^2(1+\alpha)^2}{(\alpha \theta(1+\alpha) + \alpha(2+\alpha))\Gamma(x+1)} \left( \frac{\Gamma(x+2)}{(\theta + 1)^{x+2}} + \frac{\Gamma(x+1)}{(\theta + 1)^{x+1}} - \frac{\Gamma(x+1)}{(\alpha \theta + \theta + 1)^{x+1}} \right)
\]

\[
    = \frac{\theta^2(1+\alpha)^2}{\alpha \theta(1+\alpha) + \alpha(2+\alpha)} \left( \frac{\theta + x + 2}{(\theta + 1)^{x+2}} - \frac{\alpha \theta + x + 2}{(\alpha \theta + 1)^{x+2}} \right).
\]

**Special cases**

(i) If \( \alpha \to \infty \), the \text{NPWL} distribution becomes the Poisson-Lindley distribution

(Sankaran, 1970).
(ii) If \( \alpha \to 0 \), the NPWL distribution becomes the Poisson-weighted Lindley distribution (El-Monsef & Sohsah, 2014); (Manesh, Hamzah, & Zamani, 2014) with \( \theta = 2 \).

The cumulative distribution and survival functions of the NPWL distribution are given by

\[
F(x) = P(X \leq x) = 1 - \left[ \frac{\theta^2(1+\alpha)^2}{a\theta(1+\alpha)+\alpha(2+\alpha)} \left( \frac{x+2}{\theta+1} \right)^{x+3} _2F_1\left(1, x+3; x+2; \frac{1}{\theta+1}\right)
+ \frac{1}{(\theta+1)^{x+2}} _2F_1\left(1, x+2; x+2; \frac{1}{\theta+1}\right)
- \frac{x+2}{(a\theta+\theta+1)^{x+3}} _2F_1\left(1, x+3; x+2; \frac{1}{a\theta+\theta+1}\right)
- \frac{1}{(a\theta+\theta+1)^{x+2}} _2F_1\left(1, x+2; x+2; \frac{1}{a\theta+\theta+1}\right) \right],
\]

(4)

\[
S(x) = P(X > x) = 1 - F(x),
\]

where \( _2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n} \frac{z^n}{n!} \), is the hypergeometric function; and

\[
(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = a(a+1) \ldots (a+n-1),
\]

is Pochhammer’s symbol (Johnson, Kemp, & Kotz, 2005).

Figure 1 Some pmf plots of the NPWL distribution with different values of \( \theta \) and \( \alpha \)
3. Properties

This section presents some statistical properties such as shape, factorial moments, probability generating function, moment generating function and moments of the NPWL \((\theta, \alpha)\) distribution.

3.1 Shape

Holgate’s theorem states that if \(g(\lambda)\) is the pdf of mixing distribution, which is unimodal and absolutely continuous distribution, then the pmf of the mixed Poisson distribution is unimodal (Holgate, 1970). According to this theorem, the pmf of NPWL distribution is unimodal because the pdf of NWL distribution is unimodal (Asgharzadeh, Bakouch, Nadarajah, & Sharafi, 2016).

The NPWL distribution, with the shape parameter \(\theta\) and the scale parameter \(\alpha\), is log-concave \[\frac{f(x+2)f(x)}{(f(x+1))^2} < 1\], therefore it has an increasing failure rate (Johnson, Kemp, & Kotz, 2005).

3.2 Factorial moments

**Proposition 2** Let \(X \sim \text{NPWL}(\theta, \alpha)\), then the factorial moments of \(X\) are

\[
\mu'_k = \frac{\theta^2(1+\alpha)^2}{\alpha(1+\alpha) + \alpha(2+\alpha)} \left( \frac{\Gamma(k+2)}{\theta^{k+2}} + \frac{\Gamma(k+1)}{\theta^{k+1}} - \frac{\Gamma(k+2)}{(\alpha\theta + \theta)^{k+2}} - \frac{\Gamma(k+1)}{(\alpha\theta + \theta)^{k+1}} \right).
\]

**Proof**

The \(k\)th factorial moment of mixed Poisson distribution can be written in the form
\[ \mu'_k = E[(X)_k] = \int_0^\infty \lambda^k g(\lambda) d\lambda. \]

If \( X \sim \text{NPWL}(\theta, \alpha) \), then the factorial moments of \( X \) is given by

\[
\mu'_k = \int_0^\infty \lambda^k \frac{\theta^2 (1 + \alpha)^2}{a\theta(1 + \alpha) + \alpha(2 + \alpha)} (1 + \lambda)(1 - e^{-\theta\lambda}) e^{-\theta\lambda} d\lambda \\
= \frac{\theta^2 (1 + \alpha)^2}{a\theta(1 + \alpha) + \alpha(2 + \alpha)} \int_0^\infty \left( \lambda^k e^{-\theta\lambda} + \lambda^k e^{-\theta\lambda} - \lambda^k e^{-\theta\lambda} - \lambda^k e^{-\theta\lambda} \right) d\lambda \\
= \frac{\theta^2 (1 + \alpha)^2}{a\theta(1 + \alpha) + \alpha(2 + \alpha)} \left( \frac{\Gamma(k + 2)}{\theta^{k+2}} + \frac{\Gamma(k + 1)}{\theta^{k+1}} - \frac{\Gamma(k + 2)}{(a\theta + \theta)^{k+2}} - \frac{\Gamma(k + 1)}{(a\theta + \theta)^{k+1}} \right).
\]

### 3.3 Probability generating function

**Proposition 3** Let \( X \sim \text{NPWL}(\theta, \alpha) \), then the probability generating function (pgf) of \( X \) is

\[
G(s) = \frac{\theta^2 (1 + \alpha)^2}{a\theta(1 + \alpha) + \alpha(2 + \alpha)} \left( \frac{s - 2}{(s - 1)^2} - \frac{a\theta + \theta - s + 2}{(a\theta + \theta - s + 1)^2} \right), \quad s < \theta + 1.
\]

**Proof**

The pgf of mixed Poisson distribution can be defined as

\[ G(s) = E(s^X) = \int_0^\infty e^{xs} g(\lambda) d\lambda. \]

If \( X \sim \text{NPWL}(\theta, \alpha) \), then the pgf of \( X \) is obtained by

\[
G(s) = \int_0^\infty e^{xs} \frac{\theta^2 (1 + \alpha)^2}{a\theta(1 + \alpha) + \alpha(2 + \alpha)} (1 + \lambda)(1 - e^{-\theta\lambda}) e^{-\theta\lambda} d\lambda \\
= \frac{\theta^2 (1 + \alpha)^2}{a\theta(1 + \alpha) + \alpha(2 + \alpha)} \int_0^\infty \left( \lambda e^{-(\theta-s+1)\lambda} + e^{-(\theta-s+1)\lambda} - \lambda e^{-(a\theta+\theta-s+1)\lambda} - e^{-(a\theta+\theta-s+1)\lambda} \right) d\lambda.
\]
\begin{align*}
&= \frac{\theta^2(1+\alpha)^2}{a\theta(1+\alpha)+\alpha(2+\alpha)} \left( \frac{1}{(\theta-s+1)^2} + \frac{1}{\theta-s+1} - \frac{1}{(a\theta+\theta-s+1)^2} - \frac{1}{a\theta+\theta-s+1} \right) \\
&= \frac{\theta^2(1+\alpha)^2}{a\theta(1+\alpha)+\alpha(2+\alpha)} \left( \frac{\theta-s+2}{(\theta-s+1)^2} - \frac{a\theta+\theta-s+2}{(a\theta+\theta-s+1)^2} \right), \quad s < \theta+1.
\end{align*}

### 3.4 Moment generating function

**Proposition 4** Let $X \sim \text{NPWL}(\theta, \alpha)$, then the moment generating function (mgf) of $X$ is

$$M_X(t) = \frac{\theta^2(1+\alpha)^2}{a\theta(1+\alpha)+\alpha(2+\alpha)} \left( \frac{\theta-e' + 2}{(\theta-e'+1)^2} - \frac{a\theta+\theta-e'+2}{(a\theta+\theta-e'+1)^2} \right), \quad t < \log(\theta+1).$$

The mgf is obtained trivially from the pgf as $M_X(t) = G(e')$.

### 3.5 Moments

The $k$th raw moment is obtained by taking the $k$th derivative of mgf with respect to $t$ and setting $t$ to zero.

Let $\delta_1 = \frac{\theta^2(1+\alpha)^2}{a\theta(1+\alpha)+\alpha(2+\alpha)}$, $\delta_2 = a\theta+\theta$ and $\delta_3 = \frac{\theta+2}{\theta^3} - \frac{a\theta+\theta+2}{(a\theta+\theta)^3}$, if $X \sim \text{NPWL}(\theta, \alpha)$ then the first four raw moments of $X$ are

$$\mu_1' = \delta_1 \delta_3,$$

$$\mu_2' = \delta_1 \left( \frac{\theta^2 + 4\theta + 6}{\theta^4} - \frac{\delta_2^2 + 4\delta_2 + 6}{\delta_2^5} \right),$$

$$\mu_3' = \delta_1 \left( \frac{\theta^3 + 8\theta^2 + 24\theta + 24}{\theta^5} - \frac{\delta_2^3 + 8\delta_2^2 + 24\delta_2 + 24}{\delta_2^5} \right),$$

$$\mu_4' = \delta_1 \left( \frac{\theta^4 + 16\theta^3 + 78\theta^2 + 168\theta + 120}{\theta^6} - \frac{\delta_2^4 + 16\delta_2^3 + 78\delta_2^2 + 168\delta_2 + 120}{\delta_2^6} \right).$$
The central moments of $X$ can be written in term of raw moments as

$$
\mu_k = E(X - \mu)^k = \sum_{j=0}^{k} (-1)^{k-j} \mu_j \mu^{k-j}.
$$

Thus, the first four central moments of $X$ are

$$
\mu_1 = 0,
$$

$$
\mu_2 = \delta_1 \left( \frac{\theta^2 + 4\theta + 6}{\theta^4} - \frac{\delta_2^2 + 4\delta_2 + 6}{\delta_2^4} - \delta_1 \delta_2^2 \right),
$$

$$
\mu_3 = \delta_1 \left( \frac{\theta^3 + 8\theta^2 + 24\theta + 24}{\theta^5} - \frac{\delta_2^3 + 8\delta_2^2 + 24\delta_2 + 24}{\delta_2^5} + \delta_3 \delta_2^2 \right)
$$

$$
+ \delta_3 \delta_2^2 \left( -\frac{3(\theta^2 + 4\theta + 6)}{\theta^4} + \frac{3(\delta_2^2 + 4\delta_2 + 6)}{\delta_2^4} + 2\delta_1 \delta_2^2 \right),
$$

$$
\mu_4 = \delta_1 \left( \frac{\theta^4 + 16\theta^3 + 78\theta^2 + 168\theta + 120}{\theta^6} - \frac{\delta_2^4 + 16\delta_2^3 + 78\delta_2^2 + 168\delta_2 + 120}{\delta_2^6} 
$$

$$
- 4\delta_2^3 \delta_3 \left( \frac{\theta^3 + 8\theta^2 + 24\theta + 24}{\theta^5} - \frac{\delta_2^3 + 8\delta_2^2 + 24\delta_2 + 24}{\delta_2^5} \right)
$$

$$
+ 6\delta_3^2 \delta_2 \left( \frac{\theta^2 + 4\theta + 6}{\theta^4} - \frac{\delta_2^2 + 4\delta_2 + 6}{\delta_2^4} \right) - 3(\delta_2)^4 \right).$$

The skewness, kurtosis and index of dispersion (ID) of $X \sim \text{NPWL} (\theta, \alpha)$ are

$$
\text{skewness}(X) = \frac{\mu_3}{\mu_2^{3/2}},
$$

$$
\text{kurtosis}(X) = \frac{\mu_4}{\mu_2^2},
$$

$$
\text{ID}(X) = \frac{V(X)}{E(X)} = \frac{\mu_2}{\mu_1}
$$

$$
= -\frac{\theta^4 \delta_2 \delta_3 \delta_2^2 - \theta^4 \delta_2^3 - 4\theta^4 \delta_2 - 6\theta^4 + \theta^2 \delta_2^4 + 4\theta \delta_2^3 + 6\theta^2}{\theta^2 \delta_2 \delta_2^2}.
$$
The mean, ID, skewness and kurtosis plots of the NPWL distribution are shown in Figure 2. The Figure illustrates that the mean is decreasing as $\theta$ and $\alpha$ increase. For fixed $\theta$, as $\alpha$ increases, the ID is increasing but the ID is decreasing as $\theta$ increases, for fixed $\alpha$. Moreover, we can see that the ID is greater than one, thus the NPWL distribution is overdispersed. While the skewness and kurtosis are increasing as $\theta$ and $\alpha$ increase.

Figure 2 Mean, ID, skewness and kurtosis plots of the NPWL distribution

4. Random variate generation

In this section, a NPWL random variate generation is indicated. Let $X \sim \text{Pois}(\lambda)$ and $\lambda \sim \text{NWL}(\theta, \alpha)$, then random variables from the NPWL$(\theta, \alpha)$ distribution can be generated by the following algorithm.

1. Generate $\lambda_i$ by using rnlindley function of LindleyR package (Mazucheli, Fernandes and de Oliveira, 2016) in the R programming language (R Core Team, 2018).

2. Generate $X_i$ from Pois$(\lambda_i)$, $i = 1, 2, \ldots, n$.

5. Parameter Estimation

In this section, the parameter estimates of the NPWL distribution are obtained by using maximum likelihood. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed as NPWL distribution with $\Theta = (\theta, \alpha)^T$, the parameter vector. Then the likelihood function from (3) is
\[ L(\Theta) = \prod_{i=1}^{n} \frac{\theta^2 (1 + \alpha)^2}{\alpha \theta (1 + \alpha) + \alpha (2 + \alpha)} \left( \frac{\theta + x + 2}{(\theta + 1)^{x+2}} - \frac{\theta + x + 2}{(\alpha \theta + \theta + 1)^{x+2}} \right). \]

The associated log-likelihood function can be expressed as:

\[ l(\Theta) = 2n \log(\theta) + 2n \log(1 + \alpha) - n \log(\alpha) - n \log(\theta(1 + \alpha) + 2 + \alpha) \]

\[ - \left( \sum_{i=1}^{n} x_i + 2n \right) \log(\theta + 1) - \left( \sum_{i=1}^{n} x_i + 2n \right) \log(\alpha \theta + \theta + 1) \]

\[ + \sum_{i=1}^{n} \log\left( (\theta + x_i + 2)(\alpha \theta + \theta + 1)^{x_i+2} - (\alpha \theta + \theta + x_i + 2)(\theta + 1)^{x_i+2} \right). \]

The score functions are found to be

\[ \frac{\partial l(\Theta)}{\partial \theta} = \frac{2n}{\theta} - \frac{n(1 + \alpha)}{(1 + \alpha) + 2 + \alpha} \left( \sum_{i=1}^{n} x_i + 2n \right) \frac{1}{\theta + 1} - \frac{n(1 + \alpha)}{(\alpha + 1)} \left( \sum_{i=1}^{n} x_i + 2n \right) \frac{1}{\alpha \theta + \theta + 1} \]

\[ + \sum_{i=1}^{n} \left( - (x_i + 2)(\theta + 1)^{x_i+1} (\alpha \theta + \theta + x_i + 2)(\theta + 1)^{x_i+2} - (\alpha + 1)(\theta + 1)^{x_i+2} \right) \]

\[ + \frac{(\alpha \theta + \theta + 1)^{x_i+2} (\alpha + 1)(x_i + 2)(\theta + 1)^{x_i+2} (\alpha \theta + \theta + 1)^{x_i+1} (\theta + 1)^{x_i+2} - (\alpha \theta + \theta + x_i + 2)(\theta + 1)^{x_i+2}}{(\theta + x_i + 2)(\alpha \theta + \theta + 1)^{x_i+2} (\theta + 1)^{x_i+2}}, \]

and

\[ \frac{\partial l(\Theta)}{\partial \alpha} = \frac{2n}{\alpha + 1} - \frac{n(\theta(1 + 2 \alpha) + 2(1 + \alpha))}{\alpha \theta (1 + \alpha) + 2 + \alpha} - \theta \left( \sum_{i=1}^{n} x_i + 2n \right) \frac{1}{\alpha \theta + \theta + 1} \]

\[ + \sum_{i=1}^{n} \theta (x_i + 2)(\theta + x_i + 2)(\alpha \theta + \theta + 1)^{x_i+1} (\theta + 1)^{x_i+2} - \theta (\theta + 1)^{x_i+2} \]

The score functions are set equal to zero in order to obtain the parameter estimates of the NPWL distribution. Although these equations are non-linear, the maximum likelihood estimates can be solved by the numerical methods. In this paper,
the method of moment estimators are given as the initial values for the BFGS method in
the optimx function of optimx package (Nash & Varadhan, 2011) in the R programming
language (R Core Team, 2018).

6. Applications

We consider the application of the NPWL distribution to two automobile
insurance claims data sets. These data sets are overdispersed count data with excess
zeros. The competitive distributions are as follows:

(i) The Poisson distribution (Poisson, 1837), with pmf

\[ f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad \lambda > 0 \]

(ii) The Poisson-Lindley distribution (Sankaran, 1970), with pmf

\[ f(x) = \frac{\theta^2 (x + \theta + 2)}{(\theta + 1)^{x+3}}, \quad \theta > 0 \]

(iii) The Poisson-weighted Lindley distribution (El-Monsef & Sohsah, 2014;
Manesh, Hamzah, & Zamani, 2014), with pmf

\[ f(x) = \frac{\Gamma(x + c)}{x! \Gamma(c)} \frac{\theta^{1+c}}{(\theta + c) (1 + \theta + x + c)} \frac{(1 + \theta + x + c)}{(\theta + 1)^{1+x+c}}, \quad c > 0 \text{ and } \theta > 0. \]

The mixed Poisson distributions have proportion of zeros higher than the
Poisson distribution. Thus, they have been applied to overdispersed and zero-inflated
data. Puig (2006) proposed the zero-inflation index, which is a measure for detecting
zero-inflation. If \(X\) is a non-negative integer random variable with mean and proportion
of zeros are \(\mu\) and \(p_0\), respectively. The zero-inflation index is

\[ zi(X) = 1 + \frac{\log(p_0)}{\mu}. \] (5)
If $X$ is a Poisson random variable, the zero-inflation index is equal to 0 but if $X$ is a zero-inflated, then the zero-inflation index is greater than 0. Figure 3 shows the zero-inflation index versus the index of dispersion for the PL, PWL and NPWL distributions. The zero-inflation index of the three distributions are similar. That is, the zero-inflation index is increasing as the index of dispersion increases.

**Figure 3 Zero-inflation index versus index of dispersion**

In this paper, the distribution with minimum the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the negative log-likelihood (-LL) and maximum the $p$-value based on the Anderson-Darling (AD) goodness of fit test for a discrete distribution (Choulakian, Lockhart, & Stephens, 1994) is recommended as the most appropriate distribution for fitting overdispersed count data. Moreover, the likelihood ratio (LR) test is employed to compare special cases of the NPWL distribution.

**Application 1** Lemaire (1985) presented the number of claims from third-party automobile liability in Belgium. This data set is overdispersed with the mean 0.101, the variance 0.107, the ID 1.063 and the $zi$ 0.029.

| Table 1 Distribution of number of claims from third-party automobile liability in Belgium |
Table 1 illustrates that the $p$-values based on the AD test for the PL and Poisson distributions are less than the 5% significance level, hence this data set cannot be described by the PL and Poisson distributions. The mean of all distributions are equal to the mean of the data set but the variance, ID and $z_i$ of the PWL and NPWL distributions are nearest the variance, ID and $z_i$ of the data set. However, the NPWL distribution provides the lowest negative log-likelihood, AIC, BIC and the highest $p$-value based on the AD test for a discrete distribution.

The LR statistic for testing $H_0 : \text{PWL vs } H_1 : \text{NPWL}$ is 0.36 with a $p$-value 0.549. Thus, there is no statistically significant difference between the PWL and NPWL distributions.

**Application 2** Tröbliger (1961) as cited in Klugman, Panjer and Willmot (2012) presented the count of motor vehicle insurance claims per policy in Germany during 1960. It has the mean of 0.144, the variance of 0.164, the ID of 1.136 and the $z_i$ of 0.058; therefore, this data set is overdispersed.

**Table 2 Distribution of number of automobile insurance claims in Germany 1960**

Table 2 shows that the Poisson distribution also cannot describe this data set because its $p$-value based on the AD test for a discrete distribution is less than the 5% significance level. The mean of all distributions are also equal to the mean of the data set. The variance and ID of the PL distribution and the data set are equal but the $z_i$ of the PWL distribution is closest the $z_i$ of the data set. The NPWL distribution also provides the lowest negative log-likelihood and the highest $p$-value based on the AD test for a discrete distribution.
test for a discrete distribution but the PL distribution provides the lowest AIC and BIC. However, the number of parameters of the PL distribution is less than the NPWL distribution.

In order to test $H_0 : \text{PL} \ vs \ H_1 : \text{NPWL}$, the LR statistic is 1.2 with a $p$-value 0.273. Hence, there is no statistically significant difference between the PWL and NPWL distributions. The LR statistic for testing $H_0 : \text{PWL} \ vs \ H_1 : \text{NPWL}$ is 0.42 with a $p$-value 0.517. There is also no statistically significant difference between the PWL and NPWL distributions.

The log expected values of the distributions and the log observed values are shown in Figure 4. The Figure shows that the log expected values of the NPWL and PWL distributions are close to the observed values in both data sets. However, the log expected values of the Poisson distribution are far from observed values. Thus, the Poisson distribution cannot be applied to describe the overdispersed count data.

**Figure 4 Log plots of the expected values and the observed values**

7. Conclusions

In this paper, an alternative mixed Poisson distribution, namely the new Poisson mixed weighted Lindley (NPWL) distribution, is introduced. The NPWL distribution is derived from the Poisson distribution where the parameter follows the new weighted Lindley distribution. The PL and PWL distributions are special cases of the NPWL distribution. Also, the pmf of NPWL distribution is log-concave and unimodal. Some statistical properties were studied such as shape, factorial moments, probability generating function, moment generating function and moments. Parameter estimation
was derived by the maximum likelihood estimation. Moreover, two real overdispersed count data sets were analyzed that the NPWL distribution provides a satisfactory fit in both data sets. Therefore it is an alternative distribution for modeling overdispersed count data.

Acknowledgments

The authors express their gratitude to the Department of Statistics, Faculty of Science, Kasetsart University and Rajamangala University of Technology Phra Nakhon. Further, the authors would like to thank two reviewers for valuable comments and suggestions which improved this paper.

References


Table 1 Distribution of number of claims from third-party automobile liability in Belgium

<table>
<thead>
<tr>
<th>Number of claims</th>
<th>Observed frequencies</th>
<th>Expected frequencies</th>
<th>Poisson</th>
<th>PL</th>
<th>PWL</th>
<th>NPWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>96978</td>
<td>96689.53</td>
<td>97147.15</td>
<td>96980.98</td>
<td>96981.22</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9240</td>
<td>9773.440</td>
<td>8928.854</td>
<td>9230.643</td>
<td>9231.137</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>704</td>
<td>493.953</td>
<td>816.306</td>
<td>708.754</td>
<td>706.911</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>16.643</td>
<td>74.286</td>
<td>50.019</td>
<td>50.884</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>0.421</td>
<td>6.733</td>
<td>3.372</td>
<td>3.583</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.009</td>
<td>0.608</td>
<td>0.221</td>
<td>0.250</td>
<td></td>
</tr>
</tbody>
</table>

Estimated parameters:
- $\hat{\lambda} = 0.101$
- $\hat{\theta} = 10.73$
- $\hat{\gamma} = 1.62$
- $\hat{\tau} = 14.10$
- $\hat{\alpha} = 1.835$

- Mean
  - 0.101
  - 0.101
  - 0.101
  - 0.101

- Variance
  - 0.101
  - 0.111
  - 0.107
  - 0.107

- ID
  - 1
  - 1.098
  - 1.061
  - 1.061

- $\hat{z}_i$
  - 0
  - 0.046
  - 0.029
  - 0.029

- -LL
  - 36188.25
  - 36122.53
  - 36104.11
  - 36103.93

- AIC
  - 72378.51
  - 72247.06
  - 72212.22
  - 72211.85

- BIC
  - 72388.09
  - 72256.64
  - 72231.38
  - 72231.01

- AD statistic
  - 10.301
  - 2.667
  - 0.005
  - 0.005

- p-value
  - < 0.01
  - 0.022
  - 0.991
  - 0.992
### Table 2 Distribution of number of automobile insurance claims in Germany 1960

<table>
<thead>
<tr>
<th>Number of claims</th>
<th>Observed frequencies</th>
<th>Expected frequencies</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Poisson</td>
<td>PL</td>
</tr>
<tr>
<td>0</td>
<td>20592</td>
<td>20420.94</td>
<td>20612.13</td>
</tr>
<tr>
<td>1</td>
<td>2651</td>
<td>2945.103</td>
<td>2604.386</td>
</tr>
<tr>
<td>2</td>
<td>297</td>
<td>212.371</td>
<td>326.210</td>
</tr>
<tr>
<td>3</td>
<td>41</td>
<td>10.209</td>
<td>40.562</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>0.368</td>
<td>5.013</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.011</td>
<td>0.616</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0.075</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0.009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimated parameters</th>
<th>( \hat{\lambda} = 0.144 )</th>
<th>( \hat{\theta} = 7.728 )</th>
<th>( \hat{c} = 1.103 )</th>
<th>( \hat{\phi} = 8.145 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.144</td>
<td>0.144</td>
<td>0.144</td>
<td>0.144</td>
</tr>
<tr>
<td>Variance</td>
<td>0.144</td>
<td>0.164</td>
<td>0.162</td>
<td>0.161</td>
</tr>
<tr>
<td>ID</td>
<td>1.00</td>
<td>1.136</td>
<td>1.123</td>
<td>1.121</td>
</tr>
<tr>
<td>zi</td>
<td>0</td>
<td>0.063</td>
<td>0.057</td>
<td>0.056</td>
</tr>
<tr>
<td>-LL</td>
<td>10297.84</td>
<td>10223.88</td>
<td>10223.49</td>
<td>10223.28</td>
</tr>
<tr>
<td>AIC</td>
<td>20597.69</td>
<td>20449.76</td>
<td>20450.97</td>
<td>20450.56</td>
</tr>
<tr>
<td>BIC</td>
<td>20605.75</td>
<td>20457.82</td>
<td>20467.11</td>
<td>20466.69</td>
</tr>
<tr>
<td>AD statistic</td>
<td>10.565</td>
<td>0.198</td>
<td>0.051</td>
<td>0.035</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt; 0.01</td>
<td>0.618</td>
<td>0.883</td>
<td>0.922</td>
</tr>
</tbody>
</table>
Figure 1 Some pmf plots of the NPWL distribution with different values of $\theta$ and $\alpha$

(a) ID = 10.735, skewness = 1.165, kurtosis = 5.034
(b) ID = 2.039, skewness = 1.33, kurtosis = 5.517
(c) ID = 5.932, skewness = 1.177, kurtosis = 5.064
(d) ID = 1.95, skewness = 1.48, kurtosis = 6.193
(e) ID = 1.527, skewness = 1.512, kurtosis = 6.122
(f) ID = 2.102, skewness = 1.744, kurtosis = 7.347
Figure 2 Mean, ID, skewness and kurtosis plots of the NPWL distribution
Figure 3 Zero-inflation index versus index of dispersion
Figure 4 Log plots of the expected values and the observed values