Songklanakarin J. Sci. Technol.
40 (5), 1105-1114, Sep. - Oct. 2018

http://www.sjst.psu.ac.th

Original Article

The discrete weighted exponential distribution and its applications

Panpharisa Khongthip, Mena Patummasut, and Winai Bodhisuwan*

*Department of Statistics, Faculty of Science, Kasetsart University,
Chatuchak, Bangkok, 10900 Thailand*

Received: 2 March 2017; Revised: 22 April 2017; Accepted: 3 July 2017

Abstract

We propose a new discrete distribution namely the discrete weighted exponential (dWE) distribution. It is discretization of weighted exponential distribution. Some mathematical properties are presented such as the probability mass function, the distribution function, the survival function, the probability generating function, the moment generating function, the characteristic function, the hazard function, and the reversed hazard function. For parameter estimation, we calculate the estimates based on maximum likelihood method. Furthermore, we apply the dWE distribution in real datasets. In summary, the dWE distribution could be an alternative model for count data.

Keywords: discrete analogue, count data, discrete hazard function, discrete reversed hazard function, probability generating function

1. Introduction

A probability distribution is a mathematical function that plays an important role in describing a random variable. The function leads to important characteristics, e.g., probability-generating function, mean and variance. In general, distributions are categorized by type of random variables which are discrete and continuous. The discrete distributions are applied when researchers would like to describe phenomena or events that are discrete random variables, e.g., a number of customers in bank, number of accidents at a specified intersection, or number of failures in a production process. The discrete distributions such as the binomial, Poisson, and geometric distributions are often used. If a continuous random variable of interest, the normal, exponential distributions, etc. may be focused. In fact, none of the distributions is suitable in all situations, so various researchers made much effort to develop new distributions. There are many techniques for development, such as mixture, truncated, transmuted, and compound with power series. Those techniques are uniquely focused on a continuous distribution or a discrete distribution. Some recently developed distributions are truncated Poisson (Plackett, 1953), generalized normal (Nadarajah, 2005), beta-negative binomial (Wang, 2011), transmuted Kumaraswamy (Khan, King, & Hudson, 2016), and exponentiated power Lindley geometric distributions (Alizadeh, Bagheri, Alizadeh, & Nadarajah, 2016).

For the past decade, generalization of count data distribution and lifetime distribution are of much interest. Noticeable applications of these distributions may help improve analytic results in many fields such as quality control (Skinner, Montgomery, & Runger, 2003), engineering (Haase & McPherson, 2007), insurance claim (Jean-Philippe**,** Denuit**,** & Guillen, 2008), traffic (Quddus, 2008), ecology (Lindén & Mäntyniemi, 2011), and public health (Zhou *et al.*, 2014). Particularly in reliability frameworks, the lifetime distributions such as exponential and Weibull distributions are used to describe time to failure of a device. In real situations, the lifetime data are focused on point of time or recorded in a discrete for convenience. For instance, age of people or lifetime of a device are recorded in years without a decimal. In this case, aforementioned generalization techniques may be not capable of deriving an associated distribution. Therefore, a methodology to generate a discrete distribution from a continuous distribution has recently emerged, called discretization method.

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

The most well-known type of discretization method is considered based on difference values of survival functions, which is used to create probability mass function (pmf). Nakagawa and Osaki (1975) studied the discrete failure-time distribution, the discrete Weibull distribution, and discussed a few reliability properties. Roy (1993) showed that the exponential distribution is related to the geometric distribution by using this type of procedure. Moreover, many researchers have made much effort to demonstrate discrete procedures. Roy (2003) also proposed discrete normal distribution and its properties. In addition, Kemp (2004) provided a definition of this type including reliability of discrete lifetime distribution. Some recent discrete distributions created from corresponding continuous distribution are discrete Maxwell (Krishna & Pundir, 2007), discrete Burr (Krishna & Pundir, 2009), discrete Pareto (Krishna & Pundir, 2009), discrete Lindley (Gómez-Déniz & Calderín-Ojeda, 2011), discrete generalized gamma (Chakraborty & Chakravarty, 2012), and discrete asymmetric Laplace distributions (Sangpoom & Bodhisuwan, 2016).

This paper describes a new discretized distribution created from a corresponding continuous distribution by discretization proces relying on Roy's method. Thus, a new discrete distribution will be developed from the weighted exponential (WE) distribution (Gupta & Kundu, 2009). The WE distribution is created by weighting the exponential distribution. It is a generalization of the exponential like Weibull, gamma, and generalized exponential distributions, so it can be applied in the reliability field. Moreover, the WE distribution has been extended further by researchers as the wrapped weighted exponential (Roy & Adnan, 2012), generalized weighted exponential (Kharazmi, Mahdan, & Fathi zadeh, 2015) and length biased weighted exponential distributions (Das & Kundu, 2016).

For the rest of this paper, a definition of a new discrete distribution is provided in Section 2. Some mathematical properties and a parameter estimation are shown in Section 3 and Section 4, respectively. Finally, the purposed distribution is applied to two real datasets.

2. A New Discretized Distribution

In this section, some important components for constructing a new distribution are shown such as the concept of discretization method and some features of the WE distribution. Then, a discrete weighted exponential (dWE) distribution is constructed.

2.1 The weighted exponential distribution

The WE distribution, which is a lifetime distribution extended from Azzalini's idea (Azzalini, 1985), was developed by Gupta and Kundu (2009). It is a generalized version of the exponential distribution with one additional shape parameter.

Definition 1:

Let X be distributed according to the WE random variable with shape parameter α and scale parameter λ. Its probability density function (pdf) and distribution function (Gupta & Kundu, 2009) are respectively

$$f(x)=\left(\frac{\alpha+1}{\alpha}\right)\lambda e^{-\lambda x}\left(1-e^{-\alpha\lambda x}\right),$$

and

$$(x)=\left(\frac{\alpha+1}{\alpha}\right)\left[1-e^{-\lambda x}-\left(\frac{1}{\alpha+1}\right)\left(1-e^{-(1+\alpha)\lambda x}\right)\right],$$

where $x\in(0,\infty)$ and parameters $\alpha,\ \lambda>0$.

We can see from the pdf, when $\lambda=1$ and $\alpha\to\infty$, the WE distribution converges to the exponential distribution with parameter, $\lambda=1$. It also converges to the gamma distribution with shape parameter 2 and scale parameter λ, when $\alpha\to 0$. Furthermore, it coincides with the generalized exponential distribution when shape and scale parameters are 2 and λ, repectively when $\alpha=1$ (Gupta & Kundu, 2009). Some pdf plots of the WE distribution are shown in Figure 1.

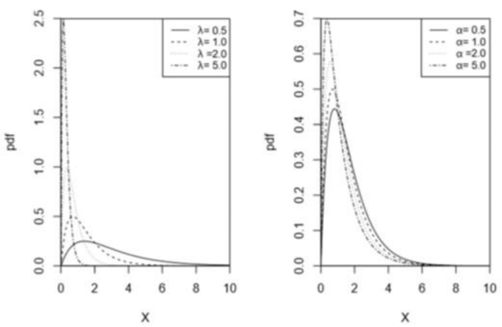

Figure 1. Some pdf plots of the WE distribution with different values of parameters λ with $\alpha=1$ (left) and of α with $\lambda=1$ (right).

2.2 A discretization method

There are many types of discretization method (Chakraborty, 2015) which obtain the pmf, e.g., the method based on the pdf, survival function, and hazard function. The pmf that obtain from the method relying on the pdf retains the structure of the pdf, while the method based on survival function maintains the form of continuous survival function, and the method based upon hazard function preserves the continuous function. Each type depends on the suitability of the characteristic properties of a continuous distribution. In this paper, difference values between survival values, $S_X(x)$ and $S_X(x+1)$, are considered. Relying upon this discretization, a full range of random variable is determined by survival functions. If X is a continuous random variable that has a survival function, $S_X(x)$, and $Y=\lfloor X\rfloor$ is the largest integer less than or equal to X. The pmf of Y is

$$\begin{aligned} p(y) &= P(Y=y) \\ &= P(\lfloor X\rfloor=y) \\ &= P(y\le x<y+1) \\ &= S_X(y)-S_X(y+1), \end{aligned} \qquad (1)$$

where $y=0,1,2,\ldots$.

2.3 The discrete weighted exponential distribution

In order to obtain a discrete version of the WE distribution, a discretization method based on survival functions of the WE distribution is employed.

Theorem 1:

If Y be a random variable of dWE distribution with parameters α and λ denoted by $Y \sim \mathrm{dWE}(\alpha,\lambda)$, then its pmf is

$$p(y)=\frac{1-e^{(\alpha+1)\lambda}+(\alpha+1)(e^{\lambda}-1)e^{(y+1)\alpha\lambda}}{\alpha e^{(\alpha+1)(y+1)\lambda}},$$

where $y \in \mathrm{N}; \mathrm{N}=\{0,1,2,\ldots\}$ and parameters $\alpha,\ \lambda>0$.

Proof:

When X follows Definition 1, its survival function is

$$S(x)=1-\left(\frac{\alpha+1}{\alpha}\right)\left[1-e^{-\lambda x}-\left(\frac{1}{\alpha+1}\right)\left(1-e^{-(\alpha+1)\lambda x}\right)\right]. \quad (2)$$

From the Equation (1), if $Y=\lfloor X \rfloor$ is the largest integer less than or equal to X then Equation (2) is plugged into Equation (1). We will obtain

$$\begin{aligned}
p(y)&=P(Y=y)\\
&=S(\lfloor X \rfloor)-S(\lfloor X+1 \rfloor)\\
&=S(y)-S(y+1)\\
&=\left\{1-\left(\frac{\alpha+1}{\alpha}\right)\left[1-e^{-\lambda y}-\left(\frac{1}{\alpha+1}\right)\left(1-e^{-(\alpha+1)\lambda y}\right)\right]\right\}\\
&\quad-\left\{1-\left(\frac{\alpha+1}{\alpha}\right)\left[1-e^{-(y+1)\lambda}-\left(\frac{1}{\alpha+1}\right)\left(1-e^{-(\alpha+1)(y+1)\lambda}\right)\right]\right\}\\
&=\left(\frac{\alpha+1}{\alpha}\right)\left[-e^{-(y+1)\lambda}+\frac{e^{-(\alpha+1)(y+1)\lambda}}{\alpha+1}+e^{-\lambda y}-\frac{e^{-(\alpha+1)\lambda y}}{\alpha+1}\right]\\
&=\left(\frac{1}{\alpha}\right)\left[\frac{(\alpha+1)(e^{\lambda}-1)e^{(y+1)\alpha\lambda}+1-e^{(\alpha+1)\lambda}}{e^{(\alpha+1)(y+1)\lambda}}\right]\\
&=\frac{1-e^{(\alpha+1)\lambda}+(\alpha+1)(e^{\lambda}-1)e^{(y+1)\alpha\lambda}}{\alpha e^{(\alpha+1)(y+1)\lambda}}.
\end{aligned}$$

Based on the following verifications we can say that $p(y)$ is said to be a pmf,

1. $p(y)\geq 0$ for all y,
2. For summation of pmf on all supporting value,

$$\begin{aligned}
\sum_{\forall y}p(y)&=\sum_{\forall y}P(Y=y)\\
&=\sum_{y=0}^{\infty}\frac{1-e^{(\alpha+1)\lambda}+(\alpha+1)(e^{\lambda}-1)e^{(y+1)\alpha\lambda}}{\alpha e^{(\alpha+1)(y+1)\lambda}}\\
&=\frac{1}{\alpha}\sum_{y=0}^{\infty}\left[e^{-(\alpha+1)(y+1)\lambda}-e^{(\alpha+1)(1-y-1)\lambda}+(\alpha+1)(e^{\lambda}-1)e^{(\alpha-1-\alpha)(y+1)\lambda}\right]\\
&=\frac{1}{\alpha}\left[-\frac{e^{-\lambda-\alpha\lambda}}{e^{-(\alpha+1)\lambda}-1}+\frac{1}{e^{-(\alpha+1)\lambda}-1}+\frac{e^{\lambda}(\alpha+1)}{e^{\lambda}-1}-\frac{\alpha+1}{e^{\lambda}-1}\right]\\
&=\frac{1}{\alpha}\left[-1+(1+\alpha)\right]\\
&=1.
\end{aligned}$$

Consequently, some probability functions will be obtained.

Theorem 2:

If Y is a dWE random variable, denoted by $Y \sim \mathrm{dWE}(\alpha,\lambda)$, then its survival function is

$$S(y)=1-\left(\frac{\alpha+1}{\alpha}\right)\left[1-\mathrm{e}^{-(y+1)\lambda}-\left(\frac{1}{1+\alpha}\right)\left(1-\mathrm{e}^{-(\alpha+1)(y+1)\lambda}\right)\right],$$

where $y \in \mathrm{N}; \mathrm{N}=\{0,1,2,\ldots\}$ and parameters $\alpha,\ \lambda>0$.

Proof:

Suppose X is a continuous random variable supporting $x \in (0,\infty)$. Similarly, $S(y)$ is defined as the survival function of a discrete random variable Y when $Y=\lfloor X \rfloor$. Thus

$$\begin{aligned} S(x+1) &= P(X \geq x+1) \\ &= P(\lfloor X \rfloor \geq \lfloor x+1 \rfloor) \\ &= P(Y \geq y+1) \\ &= 1-P(Y \leq y) \\ &= S(y). \end{aligned} \tag{3}$$

If X is a continuous random variable according to the Definition 1 and Y is a discrete random variable following the Theorem 1, then the Theorem 2 is proved.

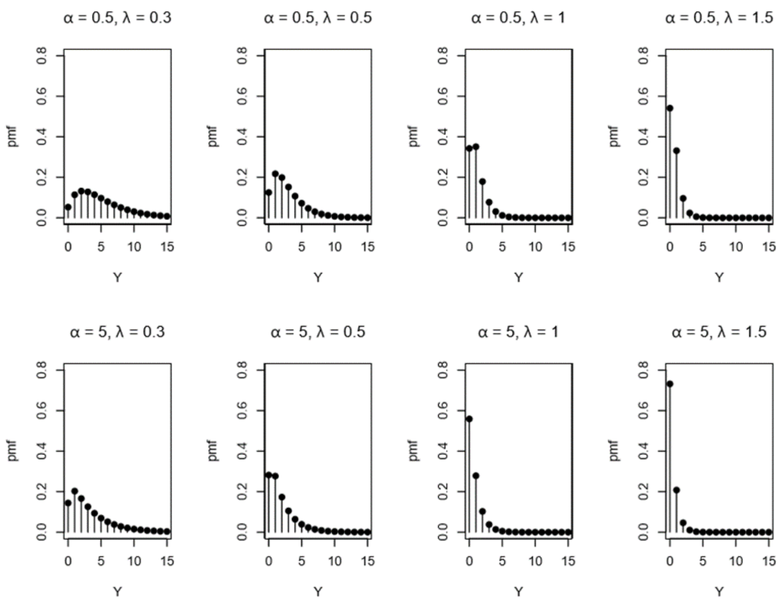

Figure 2. Some pmf plots of the dWE distribution with various values of λ and α

Figure 2 shows the pmf plots of the dWE distribution. Explicitly, the shape of distribution changes when the parameter λ changes. If λ or α increases, the distribution limit to an exponential curve. On the other hand, it has unimodal curve. Some related distributions are discussed, e.g., the dWE distribution coincides with the discrete generalized exponential (dGE) distribution with the shape parameter 2 and the scale parameter $1/\lambda$ when $\alpha=1$. It also can be reduced to the discrete exponential (dE) distribution with scale parameter $1/\lambda$ or the geometric (Geo) distribution with the parameter $\mathrm{e}^{-\lambda}$ when $\alpha \to \infty$ and the discrete gamma (dG) distribution with the shape parameter 2 and the scale parameter $1/\lambda$ when $\alpha \to 0$.

3. Some Mathematical Properties

Some mathematical properties are explored in this section, both basic properties and reliability properties, such as a probability generating function (pgf), a moment generating function (mgf), a characteristic function, a hazard function, and a reversed hazard function.

3.1 Probability generating function

The pgf plays an important role in statistical theory, because this function leads to its associated functions such as the mgf, mean, variance etc. It is defined as

$$G(s)=\sum_{y=0}^{\infty} f(y)s^{y}=E\left(s^{y}\right), \tag{4}$$

when s is a real number, and $f(y)$ is the pmf of discrete random variable y.

Theorem 3:

Suppose a random variable Y has the dWE distribution, then its pgf is given by

$$G(s)=\frac{\alpha e^{(\alpha+2)\lambda}+(s-\alpha-1)e^{(\alpha+1)\lambda}+(1-s-\alpha s)e^{\lambda}+\alpha s}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)\left(e^{\lambda}-s\right)},$$

where $y\in(0,\infty)$, s is a real number and parameters $\alpha,\ \lambda>0$.

Proof:

The pmf of the dWE distribution from the Theorem 1 is replaced in Equation (4), then the pgf is

$$\begin{aligned}
G(s)&=\sum_{y=0}^{\infty}\frac{1-e^{(\alpha+1)\lambda}+(\alpha+1)\left(e^{\lambda}-1\right)e^{(y+1)\alpha\lambda}}{\alpha e^{(\alpha+1)(y+1)\lambda}}s^{y}\\
&=\sum_{y=0}^{\infty}\frac{s^{y}}{\alpha e^{(\alpha+1)(y+1)\lambda}}-\sum_{y=0}^{\infty}\frac{s^{y}e^{(\alpha+1)\lambda}}{\alpha e^{(\alpha+1)(y+1)\lambda}}+\sum_{y=0}^{\infty}\frac{(\alpha+1)\left(e^{\lambda}-1\right)s^{y}e^{(y+1)\alpha\lambda}}{\alpha e^{(\alpha+1)(y+1)\lambda}}\\
&=\frac{1}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)}-\frac{e^{(\alpha+1)\lambda}}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)}+\frac{(\alpha+1)\left(e^{\lambda}-1\right)}{\alpha\left(e^{\lambda}-s\right)}\\
&=\frac{\alpha e^{(\alpha+2)\lambda}+(s-\alpha-1)e^{(\alpha+1)\lambda}+(1-s-\alpha s)e^{\lambda}+\alpha s}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)\left(e^{\lambda}-s\right)}.
\end{aligned}$$

The expectation and variance of Y can be simply derived from the pgf following Theorem 3. First, the first derivative of the pgf is took with respect to s, $G_Y'(s)$.

$$\begin{aligned}
\frac{d}{ds}G(s)&=\frac{d}{ds}\frac{\alpha e^{(\alpha+2)\lambda}+(s-\alpha-1)e^{(\alpha+1)\lambda}+(1-s-\alpha s)e^{\lambda}+\alpha s}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)\left(e^{\lambda}-s\right)}\\
&=\Big[\left(e^{(\alpha+1)\lambda}-s\right)\left(e^{\lambda}-s\right)\frac{d}{ds}\left(\alpha e^{(\alpha+2)\lambda}+(s-\alpha-1)e^{(\alpha+1)\lambda}+(1-s-\alpha s)e^{\lambda}+\alpha s\right)\\
&-\left(\alpha e^{(\alpha+2)\lambda}+(s-\alpha-1)e^{(\alpha+1)\lambda}+(1-s-\alpha s)e^{\lambda}+\alpha s\right)\frac{d}{ds}\left(e^{(\alpha+1)\lambda}-s\right)\left(e^{\lambda}-s\right)\Big]\\
&\times\frac{1}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)^{2}\left(e^{\lambda}-s\right)^{2}}\\
&=\Big[(\alpha+1)\left(e^{(2\alpha+3)\lambda}-e^{(2\alpha+2)\lambda}\right)-e^{(\alpha+3)\lambda}-2\alpha s e^{(\alpha+2)\lambda}+\left(-s^{2}+2\alpha s+2s\right)e^{(\alpha+1)\lambda}\\
&+e^{2\lambda}+\left(s^{2}+\alpha s^{2}-2s\right)e^{\lambda}-\alpha s^{2}\Big]\left(\frac{1}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)^{2}\left(e^{\lambda}-s\right)^{2}}\right).
\end{aligned} \tag{5}$$

By setting $s=1$ we obtain

$$\begin{aligned}
E(Y)&=\Big[(\alpha+1)\left(e^{(2\alpha+3)\lambda}-e^{(2\alpha+2)\lambda}\right)-e^{(\alpha+3)\lambda}-2\alpha e^{(\alpha+2)\lambda}+(2\alpha+1)e^{(\alpha+1)\lambda}\\
&+e^{2\lambda}+(\alpha-1)e^{\lambda}-\alpha\Big]\left(\frac{1}{\alpha\left(e^{(\alpha+1)\lambda}-1\right)^{2}\left(e^{\lambda}-1\right)^{2}}\right).
\end{aligned}$$

Similarly, the second derivative of $G(s)$, $G_Y''(s)$, is

$$\frac{d^{2}}{ds^{2}}G(s)=\frac{d^{2}}{ds^{2}}\frac{\alpha e^{(\alpha+2)\lambda}+(s-\alpha-1)e^{(\alpha+1)\lambda}+(1-s-\alpha s)e^{\lambda}+\alpha s}{\alpha\left(e^{(\alpha+1)\lambda}-s\right)\left(e^{\lambda}-s\right)}$$

$$=\Big[2\left(\mathrm{e}^{(\alpha+1)\lambda}-s\right)^2\left(\mathrm{e}^{\lambda}-s\right)^2\left(-\alpha\mathrm{e}^{(\alpha+2)\lambda}+(\alpha-s+1)\mathrm{e}^{(\alpha+1)\lambda}+(\alpha s+s-1)\mathrm{e}^{\lambda}-\alpha s\right)$$
$$+\Big((\alpha+1)\left(\mathrm{e}^{(2\alpha+3)\lambda}-\mathrm{e}^{(2\alpha+2)\lambda}\right)-\mathrm{e}^{(\alpha+3)\lambda}-2\alpha s\mathrm{e}^{(\alpha+2)\lambda}+\left(-s^2+2\alpha s+2s\right)\mathrm{e}^{(\alpha+1)\lambda}$$
$$+\mathrm{e}^{2\lambda}+\left(s^2+\alpha s^2-2s\right)\mathrm{e}^{\lambda}-\alpha s^2\Big)\Big(2\left(\mathrm{e}^{(\alpha+1)\lambda}-s\right)^2\left(\mathrm{e}^{\lambda}-s\right)+2\left(\mathrm{e}^{(\alpha+1)\lambda}-s\right)\left(\mathrm{e}^{\lambda}-s\right)^2\Big)\Big]$$
$$\times\frac{1}{\alpha\left(\mathrm{e}^{(\alpha+1)\lambda}-s\right)^4\left(\mathrm{e}^{\lambda}-s\right)^4}. \qquad (6)$$

When Equation (5) and Equation (6) are set $s=1$, the first and second derivatives of $G(s)$ are $G_Y'(1)$ and $G_Y''(1)$ respectively. Then, its variance can be obtained from the first and second derivatives of $G(s)$ is

$$\mathrm{Var}(Y)=G_Y''(1)-\left[G_Y'(1)\right]^2+G'(1). \qquad (7)$$

Consequently, the variance of Y is

$$\mathrm{Var}(Y)=\Bigg[\alpha\Big[2\left(\mathrm{e}^{(\alpha+1)\lambda}-1\right)^2\left(\mathrm{e}^{\lambda}-1\right)^2\left(-\alpha\mathrm{e}^{(\alpha+2)\lambda}+\alpha\mathrm{e}^{(\alpha+1)\lambda}+\alpha\mathrm{e}^{\lambda}-\alpha\right)$$
$$+\Big((\alpha+1)\left(\mathrm{e}^{(2\alpha+3)\lambda}-\mathrm{e}^{(2\alpha+2)\lambda}\right)-\mathrm{e}^{(\alpha+3)\lambda}-2\alpha\mathrm{e}^{(\alpha+2)\lambda}+(2\alpha+1)\mathrm{e}^{(\alpha+1)\lambda}$$
$$+\mathrm{e}^{2\lambda}+(\alpha-1)\mathrm{e}^{\lambda}-\alpha\Big)\Big(2\left(\mathrm{e}^{(\alpha+1)\lambda}-1\right)^2\left(\mathrm{e}^{\lambda}-1\right)+2\left(\mathrm{e}^{(\alpha+1)\lambda}-1\right)\left(\mathrm{e}^{\lambda}-1\right)^2\Big)\Big]$$
$$-\Big[(\alpha+1)\left(\mathrm{e}^{(2\alpha+3)\lambda}-\mathrm{e}^{(2\alpha+2)\lambda}\right)-\mathrm{e}^{(\alpha+3)\lambda}-2\alpha\mathrm{e}^{(\alpha+2)\lambda}+(2\alpha+1)\mathrm{e}^{(\alpha+1)\lambda}$$
$$+\mathrm{e}^{2\lambda}+(\alpha-1)\mathrm{e}^{\lambda}-\alpha\Big]^2+\alpha\left(\mathrm{e}^{(\alpha+1)\lambda}-1\right)^2\left(\mathrm{e}^{\lambda}-1\right)^2\Big[(\alpha+1)\left(\mathrm{e}^{(2\alpha+3)\lambda}-\mathrm{e}^{(2\alpha+2)\lambda}\right)$$
$$-\mathrm{e}^{(\alpha+3)\lambda}-2\alpha\mathrm{e}^{(\alpha+2)\lambda}+(2\alpha+1)\mathrm{e}^{(\alpha+1)\lambda}+\mathrm{e}^{2\lambda}+(\alpha-1)\mathrm{e}^{\lambda}-\alpha\Big]\Bigg]$$
$$\left[\frac{1}{\alpha\left(\mathrm{e}^{(\alpha+1)\lambda}-1\right)^4\left(\mathrm{e}^{\lambda}-1\right)^4}\right].$$

Moreover if the pgf in Theorem 3 is set $s=\mathrm{e}^t$ the mgf of Y will be obtain

$$M_Y(t)=\frac{\alpha\mathrm{e}^{(\alpha+2)\lambda}+\left(\mathrm{e}^t-\alpha-1\right)\mathrm{e}^{(\alpha+1)\lambda}+\left(1-\mathrm{e}^t-\alpha\mathrm{e}^t\right)\mathrm{e}^{\lambda}+\alpha\mathrm{e}^t}{\alpha\left(\mathrm{e}^{(\alpha+1)\lambda}-\mathrm{e}^t\right)\left(\mathrm{e}^{\lambda}-\mathrm{e}^t\right)}.$$

When s in the pgf is replace by e^{it}, the characteristic function of the dWE distribution will be obtain as

$$\phi_Y(t)=\frac{\alpha\mathrm{e}^{(\alpha+2)\lambda}+\left(\mathrm{e}^{it}-\alpha-1\right)\mathrm{e}^{(\alpha+1)\lambda}+\left(1-\mathrm{e}^{it}-\alpha\mathrm{e}^{it}\right)\mathrm{e}^{\lambda}+\alpha\mathrm{e}^{it}}{\alpha\left(\mathrm{e}^{(\alpha+1)\lambda}-\mathrm{e}^{it}\right)\left(\mathrm{e}^{\lambda}-\mathrm{e}^{it}\right)},$$

where $i=\sqrt{-1}$.

3.2 Hazard and reversed hazard functions

A hazard function, also called failure rate, is conditional probability that a unit fails in the next small time interval $x+\Delta$, given the unit has survived at x. When the unit that has survived at $x+\Delta$ fails in the next small time interval Δ, the conditional probability is called a reversed hazard function. In general, the hazard and reversed hazard functions (Gupta, 2015) are defined as

$$h(y)=h(Y=y)$$
$$=P(Y=y|Y\geq y)$$
$$=\frac{p(y)}{S(y-1)},$$
$$\lambda(y)=\lambda(Y=y)$$
$$=P(Y=y|Y\leq y)$$
$$=\frac{p(y)}{F(y)},$$

respectively, where y is non-negative integer.

Consequently, the hazard and reversed hazard functions of the dWE random variable are

$$h(y)=\frac{1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y+1)\alpha\lambda}}{\left(1-\left(\frac{\alpha+1}{\alpha}\right)\left[1-\mathrm{e}^{-\lambda y}-\left(\frac{1}{\alpha+1}\right)\left(1-\mathrm{e}^{-(\alpha+1)\lambda y}\right)\right]\right)\alpha\mathrm{e}^{(\alpha+1)(y+1)\lambda}},$$

$$\lambda(y)=\frac{1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y+1)\alpha\lambda}}{\left[1-\mathrm{e}^{(y+1)\lambda}-\left(\frac{1}{\alpha+1}\right)\left(1-\mathrm{e}^{-(\alpha+1)(y+1)\lambda}\right)\right](\alpha+1)\mathrm{e}^{(\alpha+1)(y+1)\lambda}},$$

respectively. Some plots of these function are shown in Figure 3.

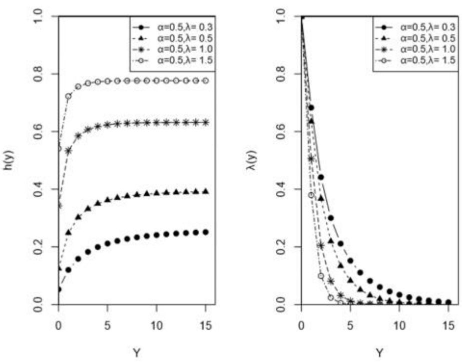

Figure 3. The plots of the hazard (left) and reversed hazard (right) functions of the dWE distribution, various the parameter values λ and α

All of the hazard functions are increasing functions. Noticeably, slope is increasing when parameter λ is increasing. Therefore, the plots of the reversed hazard functions are decreasing curve, and slope is directly dependant on parameter α .

3.3 Parameter estimation

The parameter estimation in this study will be discussed based on the maximum likelihood estimation (MLE). An associated log-likelihood function of the dWE distribution is provided by

$$l(\alpha,\lambda;y)=\sum_{i=1}^{n}\log p(y_i)$$
$$=\sum_{i=1}^{n}\log\left[\frac{1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y_i+1)\alpha\lambda}}{\alpha\mathrm{e}^{(\alpha+1)(y_i+1)\lambda}}\right]$$
$$=\sum_{i=1}^{n}\log\left(1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y_i+1)\alpha\lambda}\right)-n\log\alpha$$
$$-(\alpha+1)\lambda\sum_{i=1}^{n}y_i-n(\alpha+1)\lambda. \quad (8)$$

By taking partial derivative the Equation (8) with respected to parameters α and λ , we get

$$\frac{\partial}{\partial\alpha}l(\alpha,\lambda;y)=\frac{\partial}{\partial\alpha}\left[\sum_{i=1}^{n}\log\left(1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y_i+1)\alpha\lambda}\right)-n\log\alpha-(\alpha+1)\lambda\sum_{i=1}^{n}y_i-n(\alpha+1)\lambda\right]$$

$$=\sum_{i=1}^{n}\left[\frac{-\lambda\mathrm{e}^{(\alpha+1)\lambda}+((\alpha+1)(y_i+1)\lambda+1)\mathrm{e}^{(y_i+1)\alpha\lambda}\left(\mathrm{e}^{\lambda}-1\right)}{1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y_i+1)\alpha\lambda}}\right]-\lambda\sum_{i=1}^{n}y_i-\frac{n}{\alpha}-n\lambda. \quad (9)$$

$$\frac{\partial}{\partial\lambda}l(\alpha,\lambda;y)=\frac{\partial}{\partial\lambda}\left[\sum_{i=1}^{n}\log\left(1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y_i+1)\alpha\lambda}\right)-n\log\alpha-(\alpha+1)\lambda\sum_{i=1}^{n}y_i-n(\alpha+1)\lambda\right]$$

$$=\sum_{i=1}^{n}\left[\frac{(\alpha+1)\left(-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+y_i\alpha+1)\mathrm{e}^{(1+(y_i+1)\alpha)\lambda}-\alpha(y_i+1)\mathrm{e}^{(y_i+1)\alpha\lambda}\right)}{1-\mathrm{e}^{(\alpha+1)\lambda}+(\alpha+1)\left(\mathrm{e}^{\lambda}-1\right)\mathrm{e}^{(y_i+1)\alpha\lambda}}\right]-(\alpha+1)\sum_{i=1}^{n}y_i-n(\alpha+1). \quad (10)$$

Then the Equation (9) and Equation (10) are set equal to zero in order to obtain the maximum likelihood estimators. Nevertheless, we cannot obtain explicit expressions of the estimators, then numerical method is employed.

3.4 Applications

In this section, the dWE distribution is applied to two real datasets and compared to the discrete Lindley (dL), negative binomial (NB), Geo, dG, and dGE distributions. The first dataset is the number of deaths due to horse kicks in the Prussian army between 1875 and 1894 (Klugman, Panjer, & Willmot, 2012). The other is the prices (in ￡) of the 31 different children's wooden toys on sale in a Suffolk craft shop in April 1991 (Hand DalY, Daly, Lunn, McConway, & Ostroski, 1994). The best fits are verified with the Anderson-Darling (AD) test and other criteria including log-likelihood (LL) and the Akaike information criterion (AIC) for model selection. Furthermore, the estimated parameters are obtained by using the optim function in R language (R Core Team, 2016). Tables 1 and 2 show the MLE, LL, AD test, and AIC of the data, respectively. Figure 4 shows a comparison between real datasets and expected values of fitted distributions.

Table 1. The number of deaths due to horse kicks in the Prussian army data

Number of death	Observed values	Expected					
		dWE	dG	dGE	Geo	dL	NB
0	109	109	108	108	124	119	112
1	65	67	68	68	47	52	62
2	22	19	19	18	18	19	20
3	3	4	4	4	7	6	5
4	1	1	1	1	3	2	1
Estimated parameters		$\theta=0.0025$	$\hat{n}=2.0486$	$\hat{\alpha}=2.3953$	$\hat{p}=0.6211$	$\hat{\theta}=1.3647$	$\hat{n}=6.1209$
		$\hat{\lambda}=1.8239$	$\hat{\theta}=0.5357$	$\hat{p}=0.2253$			$\hat{p}=0.9091$
LL		-206.42	-206.41	-206.62	-213.65	-209.96	-206.53
AIC		416.83	416.81	417.24	429.30	421.93	417.06
AD statistic		0.0573	0.0646	0.0910	2.4958	1.2174	0.1128
p-value		0.9601	0.9521	0.9195	0.0322	0.1617	0.8935

Table 2. The children's wooden toy prices data

Prices	Observed values	Expected					
		dWE	dG	dGE	Geo	dL	NB
0	4	5	5	5	7	4	5
1	9	6	5	5	5	5	5
2	4	5	4	5	4	5	4
3	2	4	4	4	3	4	4
4	2	3	3	3	3	3	3
5	2	2	2	2	2	3	2
6	1	2	2	2	2	2	2
7	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1
11	2	0	0	0	0	0	0
12	1	0	0	0	0	0	0

Table 2. (Continued)

Estimated Parameters	$\hat{\theta}=8.8355$ $\hat{\lambda}=0.2623$	$\hat{n}=1.2470$ $\hat{\theta}=3.3573$	$\hat{\alpha}=1.2795$ $\hat{p}=0.7582$	$\hat{p}=0.2123$	$\hat{\theta}=0.4072$	$\hat{n}=1.2931$ $\hat{p}=0.2587$
LL	-74.7811	-75.1343	-75.1388	-75.4858	-75.6413	-75.2366
AIC	153.5623	154.2686	154.2777	152.9715	153.2825	154.4733
AD statistic	0.3446	0.4053	0.4065	0.3719	0.7248	0.3996
p-value	0.7873	0.7215	0.7209	0.7256	0.4635	0.7231

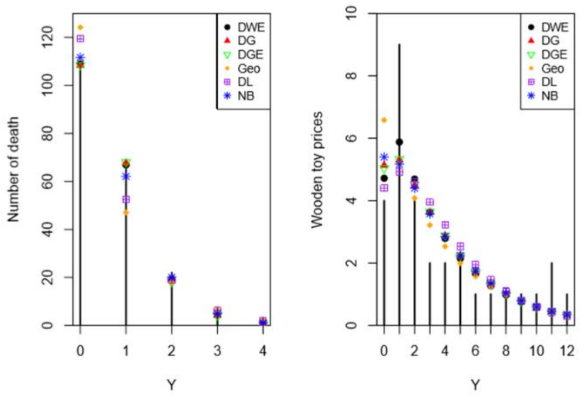

Figure 4. Real data and expected values of fitted distributions with various discrete distributions

For primary consideration of appropriatness of the distributions, the LL and AIC values are concerned. From Table 1, although, the dG distribution has the biggest LL value and the smallest AIC value, the dWE distribution is the most appropriate based on AD goodness of fit test for the number of deaths due to horse kicks in the Prussian army between 1875 and 1894 data. In the same manner, the dWE distribution is more appropriate than the others based on LL and AD goodness of fit test for the children's wooden toy prices data in Table 2.

4. Conclusions

In this study, the new count data distribution named the dWE distribution is developed by discretization of the WE distribution. The pmf of the dWE distribution is derived, some mathematical properties are discussed, and reliability functions are presented. In addition, the dWE distribution is applied to two real datasets and compared to the others based on MLE. We choose the dG, dGE, and Geo distributions for comparison because they are related to dWE distribution. The dL and NB distributions have decreasing and unimodal shapes like the dWE distribution. The result is that the dWE fit better than the others. Thus, the dWE distribution can be an alternative count data distribution.

Acknowledgements

The first author would like to thank the Science Achievement Scholarship of Thailand (SAST).

References

Alizadeh, M., Bagheri, S. F., Alizadeh, M., & Nadarajah, S. (2016). A new four-parameter lifetime distribution. *Journal of Applied Statistics.* doi:10.1080/02664763.2016.1182137

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics, 12*(2), 171–178.

Chakraborty, S., & Chakravarty, D. (2012). Discrete gamma distributions: properties and parameter estimation. *Communication in Statistics-Theory and Methods, 40*(18), 3301–3324.

Chakroborty, S. (2015). Generating discrete analogues of continuous probability distributions - A survey of methods and constructions. *Journal of Statistical Distributions and Applications, 2*(6). doi:10.1186/s40488-015-0028-6

Das, S., & Kundu, D. (2016). On weighted exponential distribution and its length biased version. *The Indian Society for Probability and Statistics, 17*, 57–77.

Gómez-Déniz, E., & Calderín-Ojeda, E. (2011). The discrete lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation, 81*(11), 1405–1416.

Gupta, P. L. (2015). Properties of reliability functions of discrete distributions. *Communications in Statistics, 44*(19), 4114–4131.

Gupta, R. D., & Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics, 43*(6), 621–634.

Haase, G. S., & McPherson, J. W. (2007). Modeling of interconnect dielectric lifetime under stress conditions and new extrapolation methodologies for time-dependent dielectric breakdown. *2007 IEEE International Reliability Physics Symposium Proceedings 45th Annual*, 390-398. doi:10.1109/RELPHY.2007.369921

Hand, D. J., Daly, F., Lunn, A. D., McConway, K. J., & Ostrowski, E. (1994). *A handbook of small data sets.* London, England: Chapman and Hall.

Jean-Philippe, B., Denuit, M., & Guillen, M. (2008). Models of insurance claim counts with time dependence based on generalization of poisson and negative binomial distributions. *Variance, 2*(1), 135-162.

Kemp, A. W. (2004). Classes of discrete lifetime distributions. *Communications in Statistics - Theory and Methods, 33*(12), 3069–3093.

Khan, M. S., King, R., & Hudson, I. L. (2016). Transmuted kumaraswamy distribution. *Statistics in Transition New Series, 17*(3), 183–210.

Kharazmi, O., Mahdavi, A., & Fathizadeh, M. (2015). Generalized weighted exponential distribution. *Communications in Statistics - Simulation and Computation, 44*(6), 1557–1569.

Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). *Loss models: From data to decisions* (3th Ed.). Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley and Sons.

Krishna, H., & Pundir, P. S. (2007). Discrete maxwell distribution. *Interstat.* Retrived from http://interstat.statjournals.net/YEAR/2007/articles/0711003.pdf.

Krishna, H., & Pundir, P. S. (2009). Discrete burr and discrete pareto distributions. *Statistical Methodology, 6*(2), 177–188.

Lindén, A., & Mäntyniemi, S. (2011). Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology, 92*(7), 1414-1421.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics, 32*(7), 685–694.

Nakagawa, T., & Osaki, S. (1975). The discrete weibull distribution. *IEEE Trnsactions on Reliability, 24*(5), 300–301.

Plackett, R. L. (1953). The truncated poisson distribution. *Biometrics, 9*(4), 485–488.

Quddus, M. A. (2008). Time series count model: An empirical application to traffic acidents. *Accident Analysis and Prevention, 40*, 1732-1741.

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Roy, D. (1993). Reliability measures in the discrete bivariate set up and related characterization results for a bivariate geometric distribution. *Journal of Multivariate Analysis, 46*(2), 362–373.

Roy, D. (2003). The discrete normal distribution. *Communications in statistics - Theory and Methods, 32*(10), 1871–1883.

Roy, S., & Adnan, M. A. S. (2012). Wrapped weighted exponential distributions. *Statistics and Probability Letters, 82*(1), 77–83.

Sangpoom, S., & Bodhisuwan, W. (2016). The discrete asymmetric laplace distribution. *Journal of Statistical Theory and Practice, 10*(1), 73–86.

Skinner, K. R., Montgomery, D. C., & Runger, G. C. (2003). Process monitoring for multiple count data using generalized linear model-based control charts. *International Journal of Production Research, 41*(5), 1167-1180.

Wang, Z. (2011). One mixed negative binomial distribution with application. *Journal of Statistical Planning and Inference, 141*(3), 1153–1160.

Zhou, H., Siegel, P. Z., Barile, J., Njai, R. S., Thompson, W. W., Kent, C., & Liao, Y. (2014). Models for count data with an application to healthy days measures: Are you driving in screws with a hammer? *Preventing Chronic Disease, 11*(E50). doi:10.5888/pcd11.130252