A New Generalization of Weighted Geometric Distribution and its Properties

H. Najarzadegan

Department of Statistics, University of Isfahan
Isfahan, Iran
hnajarzadegan@yahoo.com

M. H. Alamatsaz*

Department of Statistics, University of Isfahan
Isfahan, Iran
alamatho@sci.ui.ac.ir

Discrete distributions are widely used to model lifetime for count data. In this paper we introduce a new generalization of weighted geometric (GWG) distribution with the weight function $w(x) = (1 - p^{\alpha x})^{\beta}$ whose special case proposes a discrete generalized weighted exponential distribution. Further, it is observed that GWG distribution can be produced through a selection model. GWG distribution can be viewed as the generalization of a weighted geometric distribution, the discrete generalized exponential distribution and the classic geometric distribution. We shall study several distributional and structural properties of our distribution such as its shape properties, unimodality, infinite divisibility, moments probability weighted moments, stochastic orderings, order statistics, entropies and stress-strength reliability function. We shall also present some related novel characterizations. Finally, estimation of the parameters are investigated and applicability of the model is examined using a real data set .

Keywords: Entropy, Moment, Moment generating function, Stress-strength reliability function, Stochastic orders, Order statistics, MLE, Characterization.

2000 Mathematics Subject Classification: 60E05, 62E15

1. Introduction

In the last decades, several traditional discrete distributions such as geometric, negative binomial, Poisson, etc. have been used to fit all types of discrete data. However, such classical disctere distributions are not much flexible to fit many types of discrete data. Thus, recently several methods have been proposed to construct new discrete distributions with higher flexibility to model such data. Further, in real situations, sometimes we can not measure the life length of a device on a continuous

---

*Corresponding Author.

scale. Thus, there is a real need to find more discrete distributions to cover various types of count data.

One of the most common methods for constructing new discrete distributions is to discretize known continuous distributions. Infact, there are two general techniques for generating new discrete distributions using a given baseline continuous distribution. By these techniques, we define a probability mass functions (pmf) by

$$P(Y = y) = \frac{f(y)}{\sum_{j=1}^{\infty} f(j)}, \qquad y = 1, 2, ..., \tag{1.1}$$

or

$$P(Y = y) = \bar{F}(y) - \bar{F}(y+1), \qquad y = 0, 1, ..., \tag{1.2}$$

using any probability density function (pdf) f or survival function $\bar{F}$ of a baseline absolutely continous distribution. For instance, the discrete normal distribution was introduced by Lisman and van Zuylen (1972). Similarly, Laplace, skew-Laplace, discrete additive Weibull, exponentiated discrete Weibull, discrete beta-exponential and alpha-skew-Laplace distributions on the lattice of integers were proposed by Inusah and Kozubowski (2006), Kozubowski and Inusah (2006), Bebbington (2012), Nekoukhou and Bidram (2015), Nekoukhou et al. (2015) and Harandi and Alamatsaz (2015), respectively.

Another method to construct a new discrete distribution is to define weighted version of a baseline distribution, as proposed by Patil and Rao (1977, 1978). For example Bhati and Joshi (2015) proposed the weighted geometric distribution using this method.

In this paper, we propose a new generalization of the weighted geometric (GWG) distribution by implementing Patil and Rao's (1977, 1978) method on geometric distribution with the weight function $w(x) = (1 - p^{\alpha x})^{\beta}$. It is interesting to see that our GWG distribution proposes a discrete version of the generalized weighted exponential of Kharazmi et al. (2015) as its special case.

Formerly, many attempts have been made to construct some generalizations of the geometric distribution in the literature such as those of Philippou et al. (1983), Tripathi et al. (1987), Gómez-Déniz (2010), Nekoukhou et al. (2011), Akinsete et al. (2014) , Bhati and Joshi (2015), Bidram et al. (2015), Nekoikhou and Bidram (2016) and Akdoğan et al. (2016).

Let us recall that a random variable Y has a weighted exponential (WE) distribution, denoted by $WE(\alpha, \lambda)$, if the pdf of Y is

$$f_Y(y; \alpha, \lambda) = \frac{\alpha + 1}{\alpha} \lambda e^{-\lambda y} (1 - e^{-\alpha \lambda y}), \qquad y > 0, \tag{1.3}$$

where $\alpha > 0$ and $\lambda > 0$ are the corresponding shape and scale parameters, respectively. Since WE distribution is of interest in several applications, especially in lifetime modeling and reliability, several authors have attempted to generalize this distribution to cover more applications. For instance, Shakhatreh (2012) generalized exponential distribution with a two-parameter weighted function and a regular exponential distribution; $TWE(\alpha_1, \alpha_2, \lambda)$. A special case of this model is the two parameter generalized exponential distribution when $\alpha_1 = \alpha_2 = \alpha$ with pdf

$$f_Y(y; \alpha, \alpha, \lambda) = \frac{(\alpha + 1)(1 + 2\alpha)}{2\alpha^2} \lambda e^{-\lambda y} (1 - e^{-\alpha \lambda y})^2, \quad y > 0, \alpha > 0, \lambda > 0 \tag{1.4}$$

More recently, Kharazmi et al. (2015) introduced a new class of generalized weighted exponential $(GWE(\alpha,\lambda,n))$ distributions with pdf

$$f_Y(y;\alpha,\lambda,n) = \frac{\alpha}{B(\frac{1}{\alpha},n+1)}\lambda e^{-\lambda y}(1-e^{-\lambda\alpha y})^n, \qquad y > 0, \alpha > 0, \lambda > 0, \qquad (1.5)$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ and $n \in N = \{1,2,...\}$. GWE distribution contains both WE and TWE distributions, as its sub models.

The remainder of the paper is organized as follows. In Section 2, we define the new three parameters distribution by using Patil and Rao's (1977, 1978) method on geometric distribution which we call it generalized weighted geometric (GWG) distribution. Then, we reveal that a discrete generalized weighted exponential (DGWE) distribution is obtained as its special case. We also propose a selection model to construct a DGWE distribution. In Section 3, we discuss its several mathematical aspects such as shape properties, moment generating function, different types of moments, failure rate function, mean residual life time, stochastic orders, order statistics, entropies and stress-strength reliability function. Some characterizations for special cases of the class of GWG distributions are presented in Section 4. In this section, we also reveal that GWG distribution can be viewed as a generalization of a weighted geometric (WG) distribution proposed by Bhati and Joshi (2015), the discrete gamma$(2,p)$ and the classic geometric distributions. Maximum likelihood estimation of the model parameters and the observed information matrix are discussed in Section 5. Finally, in Section 6, we show GWG distribution's potential in application by fitting it to rank frequencies of graphemes in a Slavic language of Makčutek (2008) and compare it with some rival distributions.

2. Definition and Basic Properties

In this section, we introduce the generalized weighted geometric (GWG) distribution by implementing Patil and Rao's (1977, 1978) method on geometric distribution. Then, we obtain the discrete generalized weighted exponential (DGWE) as a special case of GWG distribution. We also construct this distribution alternatively by a selection model.

Definition 2.1. Let $X_1 \sim G_1(1-p)$ be a random variable of the classic geometric distribution $G_1(1-p)$ with probability mass function (pmf)

$$p_1(x) = P(X=x) = (1-p)p^{x-1}, \qquad x = 1,2,.... \qquad (2.1)$$

Then we say that its weighted r.v. X with the weight function $w(x) = (1-p^{\alpha x})^\beta$ has a GWG distribution with parameters $\alpha > 0, 0 < p < 1$ and $\beta > 0$, i.e., when X has the following pmf

$$p(x) = \frac{w(x)p_1(x)}{E(w(X_1))} = \frac{(1-p^{\alpha x})^\beta(1-p)p^{x-1}}{E((1-p^{\alpha X_1})^\beta)} = \frac{(1-p^{\alpha x})^\beta(1-p)p^{x-1}}{\sum_{j=0}^\infty \binom{\beta}{j}(-1)^j\frac{(1-p)p^{\alpha j}}{1-p^{\alpha j+1}}}. \qquad (2.2)$$

Clearly, by Equation (2.2) we have

$$p(x) = A^{-1}p^{x-1}(1-p^{\alpha x})^\beta \qquad x = 1,2,... \qquad (2.3)$$

where $A = \sum_{j=0}^\infty \binom{\beta}{j}(-1)^j\frac{p^{\alpha j}}{1-p^{\alpha j+1}}$ and $\alpha > 0, 0 < p < 1, \beta > 0$. We shall denote such a distribution by $GWG(\alpha,\beta,p)$.

Now, by inserting the generalized weighted exponential density (1.5) into (1.1) with $p = e^{-\lambda}$, we obtain a discrete generalized weighted exponential DGWE(α, n, p) distribution with parameters $\alpha > 0$, $0 < p < 1$ and $n \in N$ if

$$p(x) = P(X = x) = \frac{p^{x-1}(1 - p^{\alpha x})^n}{\sum_{j=0}^{n} \binom{n}{j}(-1)^j \frac{p^{\alpha j}}{1 - p^{\alpha j+1}}}, \qquad x \in N = \{1, 2, ...\}. \tag{2.4}$$

That is, GWG(α, β, p) coinsides with the DGWE(α, n, p) when $\beta = n$. Using Arellano et al. (2006)'s method of selection model, we can also produce a GWG (a DGWE) distribution as follows.

Theorem 2.1. *Suposse that $X, X_1, ..., X_n$ is a random sample from the classic geometric distribution $G_1(1 - p)$ with pmf*

$$p_1(x) = P(X = x) = (1 - p)p^{x-1}, \qquad x = 1, 2, ... \tag{2.5}$$

Then, for any $\beta = n$ and $\alpha = m$; $n, m \in N$; we have

$$X \mid max(X_1, X_2, ..., X_n) \leq mX \sim GWG(m, n, p)$$

Proof. Since $X, X_1, X_2, ..., X_n$ are independant discrete random variables and $m \in N$, we obtain

$$P\big(X = x \mid max(X_1, X_2, ..., X_n) \leq mX\big) = \frac{P\big(X = x, max(X_1, X_2, .., X_n \leq mx)\big)}{P\big(max(X_1, X_2, .., X_n \leq mX)\big)},$$
$$= \frac{(1 - p)p^{x-1}(1 - p^{mx})^n}{P\big(max(X_1, X_2, .., X_n \leq mX)\big)}, \tag{2.6}$$

where

$$P\big(max(X_1, X_2, ..., X_n) \leq mX\big) = \sum_{x=1}^{\infty} P(X = x)P\big(max(X_1, X_2, ..., X_n) \leq mx\big),$$
$$= (1 - p) \sum_{j=0}^{n} \binom{n}{j}(-1)^j \frac{p^{mj}}{1 - p^{mj+1}}, \tag{2.7}$$

which implies the result.
□

Remark 2.1. GWG(α, β, p) distribution reduces to the DGE(γ, p) distribution proposed by Nekoukhou et al. (2011) when $\alpha = 1$, with $\beta = \gamma - 1$.

Theorem 2.2. *Let $X \sim GWG(\alpha, \beta, p)$. The cumulative distribution function (cdf) of X is given by*

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < 1 \\ 1 - A^{-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} \frac{p^{(\alpha j+1)[x] + \alpha j}}{1 - p^{(\alpha j+1)}} & x \geq 1, \end{cases} \tag{2.8}$$

Proof. For any $x \geq 1$, we can write

$$F(x) = P(X \leq x) = A^{-1} \sum_{t=1}^{[x]} p^{t-1} (1 - p^{\alpha t})^{\beta},$$

$$= A^{-1} \sum_{t=1}^{[x]} p^{t-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} p^{\alpha t j},$$

$$= 1 - A^{-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} \frac{p^{(\alpha j + 1)[x] + \alpha j}}{1 - p^{(\alpha j + 1)}}.$$

Thus, we have the result. □

It, then, follows that the survival function of a $GWG(\alpha, \beta, p)$ random variable X is given by

$$\bar{F}(x) = 1 - F(x) = A^{-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} \frac{p^{(\alpha j + 1)[x] + \alpha j}}{1 - p^{(\alpha j + 1)}}, \qquad (2.9)$$

for $x > 1$.

3. Distributional properties

In this section, we provide several mathematical properties of a GWG distribution. Here, we describe shape, moment generating function and moment properties and study its failure rate function, mean residual life time, stochastic orderings, order statistics, entropy and stress-stregnth reliability function.

3.1. *Shape properties*

Figure 1 illustrates possible shapes of the pmf of a GWG distribution for several selected parameter values. As we see, it seems that the distribution is unimodal. In what follows, we show that indeed more strongly the distribution is infact strongly unimodal and thus it is unimodal.

Theorem 3.1. *The pmf of a $GWG(\alpha, \beta, p)$ is strongly unimodal for all $\alpha > 0$, $0 < p < 1$ and $\beta > 0$.*

Proof. Clearly, $p(x)$, pmf of $GWG(\alpha, \beta, p)$, satisfies

$$\frac{p(x)}{p(x-1)} = \frac{p(1 - p^{\alpha x})^{\beta}}{(1 - p^{\alpha(x-1)})^{\beta}}, \qquad \alpha > 0, \quad \beta > 0, \quad 0 < p < 1,$$

which is a non-increasing function of x. Thus, by a result of keilson and Gerber (1971), $GWG(\alpha, \beta, p)$ is logconcave and, consequently, it is strongly unimodal and thus unimodal for all $\alpha > 0$, $\beta > 0$ and $0 < p < 1$. □

It is interesting to note that by a closer look at Figure 1, we can see that the mode of the distribution moves to the left as α increases for fixed β and p. However, the mode moves to the right when either p increases and α and β are fixed or when β increases and α and p are fixed.
It is also noted that, however, the distribution in question is not infinitely divisible. This is simply noticed by property (1.2) in Steutel and van Harn (2004) and the fact that $p(0) = 0$. Indeed, from here we may also conclude that a GWG model can not be neither discrete self-decomposable nor discrete stable (see, Steutel and var Harn (2004)).
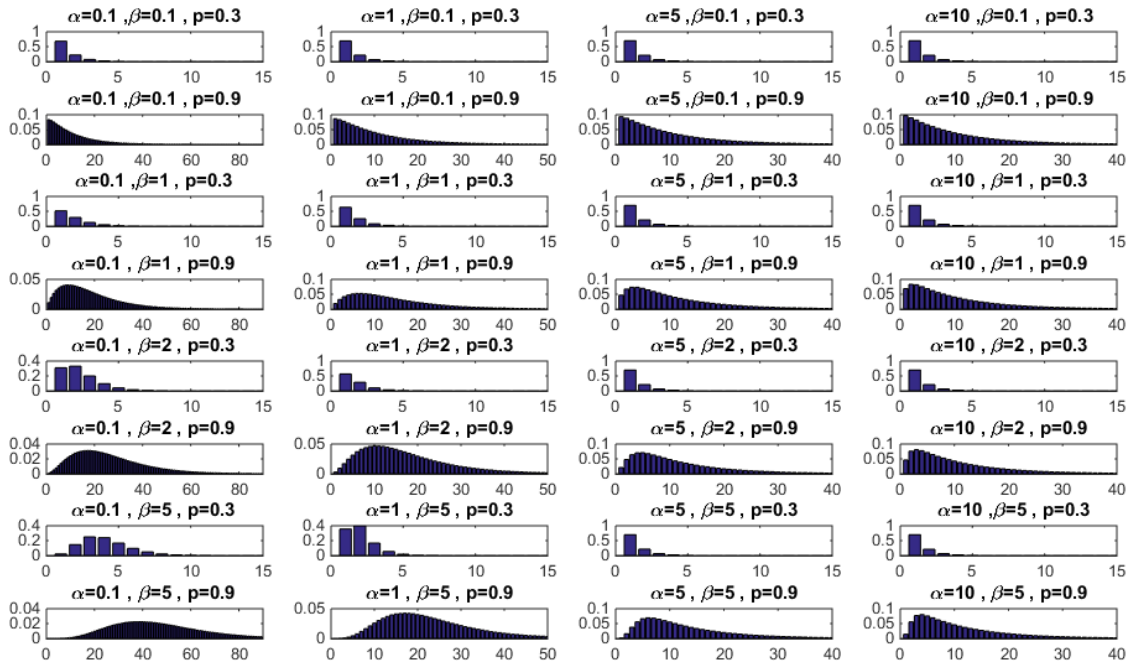
Fig. 1. Plots of pmf of $GWG(\alpha, \beta, p)$ distributions for selected values of α, β and p.

3.2. *Moment generating function and moments*

Here, we obtain moment generating function and rth-order moments of a $GWG(\alpha, \beta, p)$ distribution. Let $M_X(t) = E(e^{tX})$ be the moment generating function (mgf) of a GWG random variable X. Then, we have the following:

Theorem 3.2. *Let* $X \sim GWG(\alpha, \beta, p)$. *Then, mgf of X is given by*

$$M_X(t) = A^{-1} \sum_{j=0}^{\infty} (-1)^{-1} \binom{\beta}{j} \frac{e^t p^{\alpha j}}{1 - e^t p^{\alpha j+1}}, \qquad t < \log \frac{1}{p}. \qquad (3.1)$$

Proof.

$$M_X(t) = A^{-1} \sum_{x=1}^{\infty} e^{xt} p^{x-1} (1 - p^{\alpha x})^{\beta},$$

$$= A^{-1} \sum_{x=1}^{\infty} e^{xt} p^{x-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} p^{\alpha j x},$$

$$= A^{-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} \frac{e^t p^{\alpha j}}{1 - e^t p^{\alpha j+1}}, \qquad t < \log \frac{1}{p}.$$

$\square$

Theorem 3.3. *If $X \sim GWG(\alpha, \beta, p)$ and $\mu_r = E(X^r)$ are rth-order moments about the origin, then we have*

$$\mu_r = A^{-1} \sum_{j=0}^{\infty} \sum_{x=1}^{r} (-1)^j \binom{\beta}{j} \frac{x! p^{(\alpha j+1)x-1}}{(1-p^{\alpha j+1})^{x+1}} S(r,x), \qquad (3.2)$$

where

$$S(r,x) = \frac{1}{x!} \sum_{i=0}^{x-1} (-1)^i \binom{x}{i} (x-i)^r \qquad (3.3)$$

is the Stirling number of the second kind.

Proof. By using the known equation

$$\sum_{x=1}^{\infty} x^r p^x = \sum_{x=1}^{r} S(r,x) \frac{x! p^x}{(1-p)^{1+x}}, \qquad (3.4)$$

we obtain

$$
\begin{aligned}
\mu_r &= A^{-1} \sum_{x=1}^{\infty} x^r p^{x-1} (1-p^{\alpha x})^{\beta}, \\
&= A^{-1} \sum_{x=1}^{\infty} x^r p^{x-1} \sum_{j=0}^{\infty} (-1)^j \binom{\beta}{j} p^{\alpha jx}, \\
&= A^{-1} \sum_{j=0}^{\infty} \sum_{x=1}^{r} (-1)^j \binom{\beta}{j} \frac{x! p^{(\alpha j+1)x-1}}{(1-p^{\alpha j+1})^{x+1}} S(r,x).
\end{aligned}
$$

As required. $\square$

Hence, the corresponding variance $\sigma^2 = \mu_2 - \mu_1^2$, skewness $\gamma_1 = (\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3)/(\mu_2 - \mu_1^2)^{\frac{3}{2}}$ and kurtosis $\gamma_2 = (\mu_4 - 4\mu_1\mu_3 + 6\mu_2\mu_1^2 - 3\mu_1^4)/(\mu_2 - \mu_1^2)^2$ of $GWG(\alpha, \beta, p)$ can be easily obtained using Eq. (3.4) in explicit forms. Figures 2 and 3 plot the mean, variance, skewness and kurtosis of a $GWG(\alpha, \beta, p)$ distribution for different selected values of α, β and p. As we see in Figure 2, the mean and variance are decreasing as α increases when β and p are fixed. Further, the mean and variance are increasing either when p increases and α and β are fixed or when β increases and α and p are fixed.

On the other hand, in Figure 3, the skewness and kurtosis are increasing either as α increases and β and p are fixed or when p decreases and α and β are fixed. Further the skewness and kurtosis are decreasing as β increases and α and p are fixed.

Now, we obtain the probability weighted moments (PWMs) of a GWG distribution. They are defined when the ordinary moments of the random variable exist. The PWMs can generally be used for estimating parameters of a distribution whose inverse form cannot be expressed explicitly. They may also be used as starting values for maximum likelihood estimates. The (k, r)-th PWM of a
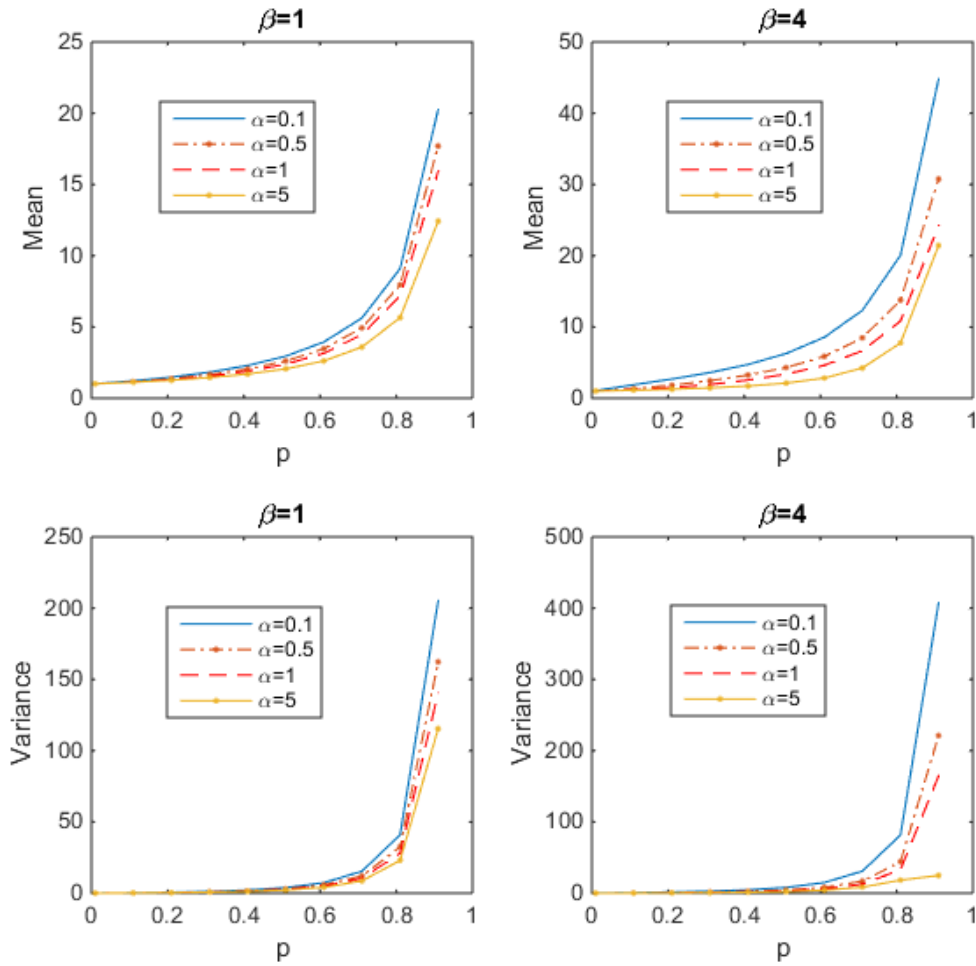
Fig. 2. Plots of mean, variance of a $GWG(\alpha, \beta, p)$ distribution for different values of α, β and p.

discrete random variable X with pmf $p(x)$ is formally defined by

$$\tau_{k,r} = E[X^k F(X)^r] = \sum_{x=1}^{\infty} x^k F(x)^r p(x). \tag{3.5}$$

Thus, by using the identity

$$\left(\sum_{j=0}^{\infty} a_j x^j\right)^t = \sum_{j=0}^{\infty} d_{t,j} x^j, \tag{3.6}$$

where $d_{t,j} = (ja_0)^{-1} \sum_{m=1}^{j} [m(t+1) - j] a_m d_{t,j-m}$ and $d_{t,0} = a_0^t$ which is defined by Gradshteyn and Ryzhik (2007) for a positive integer t and Equation (3.5), we have
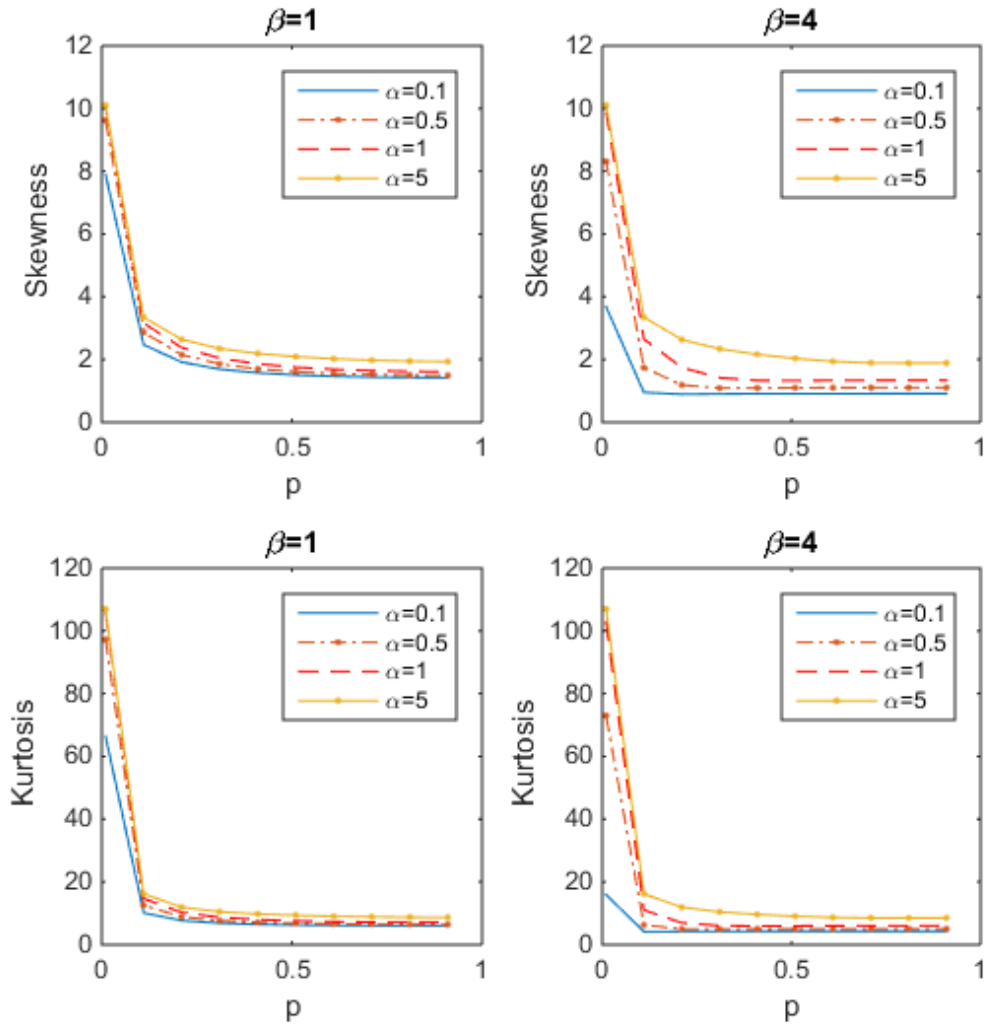
Fig. 3. Plots of skewness and kurtosis of a $GWG(\alpha, \beta, p)$ distribution for different values of α, β and p.

$$
\begin{aligned}
\tau_{k,r} &= \sum_{x=1}^{\infty} x^k \left\{ 1 - A^{-1} \sum_{j=0}^{\infty} A_j p^{(\alpha j+1)x} \right\}^r A^{-1} p^{x-1} (1 - p^{\alpha x})^{\beta}, \\
&= \sum_{x=1}^{\infty} \sum_{t=0}^{r} A^{-t-1} \binom{r}{t} (-1)^t x^k \left\{ \sum_{j=0}^{\infty} A_j p^{\alpha jx} \right\}^t p^{xt+x-1} (1 - p^{\alpha x})^{\beta}, \\
&= \sum_{x=1}^{\infty} \sum_{t=0}^{r} \sum_{l,j=0}^{\infty} d_{t,j} A^{-t-1} \binom{r}{t} \binom{\beta}{l} (-1)^{t+l} x^k p^{j\alpha x + \alpha x l + xt + x - 1}, \\
&= \sum_{l,j=0}^{\infty} \sum_{x=1}^{k} \sum_{t=0}^{r} d_{t,j} A^{-t-1} \binom{r}{t} \binom{\beta}{l} (-1)^{t+l} \frac{S(k,x) x! \, p^{\alpha x(j+l)+xt+x-1}}{(1 - p^{\alpha(j+l)+t+1})^{x+1}},
\end{aligned}
\tag{3.7}
$$

where $A_j = (-1)^j \binom{\beta}{j} \frac{p^{\alpha j}}{1-p^{(\alpha j+1)}}$, $A = \sum_{j=0}^{\infty} A_j$, $d_{t,j} = \frac{1-p}{j} \sum_{m=1}^{j} [m(t+1)-j]A_m$ and $d_{t,0} = \frac{1}{(1-p)^2}$.

3.3. *Failure rate function and mean residual lifetime*

Now we consider failure rate and mean residual life functions of a *GWG* distribution. The failure rate function is one of the most popular concepts in reliability studies. If $r(t) = p(x)/\bar{F}(x)$ is the failure rate function of a *GWG* random variable X, then based on (2.1) and (2.8) we have

$$r(x) = \frac{(1-p^{\alpha x})^{\beta}}{\sum_{j=0}^{\infty} A_j p^{\alpha jx+1}}, \qquad x = 1, 2, ... \qquad (3.8)$$

Remark 3.1. By Theorem 3.1., since $p(x)$ is log-concave, the failure rate function $r(x)$ is a non-decreasing function of x.

Figures 4 and 5 illustrate the failure rate functions of a *GWG* distribution for different values of α, β and p. A popular property of a lifetime distribution is its mean residual lifetime. For a *GWG* distribution it is not difficult to see that

$$Pr[X - x > y | X > x] = \frac{\sum_{j=0}^{\infty} A_j p^{(\alpha j+1)(x+y)}}{\sum_{j=0}^{\infty} A_j p^{(\alpha j+1)x}}, \qquad (3.9)$$

for $x = 1, 2, ...$ and $y = 0, 1, ...$. Thus, from Eq. (3.9), the mean residual lifetime of X is obtained as

$$m(x) = E(X - x | X > x) = \frac{\sum_{j=0}^{\infty} \frac{A_j p^{\alpha jx}}{1-p^{(\alpha j+1)}}}{\sum_{j=0}^{\infty} A_j p^{\alpha jx}}, \qquad (3.10)$$
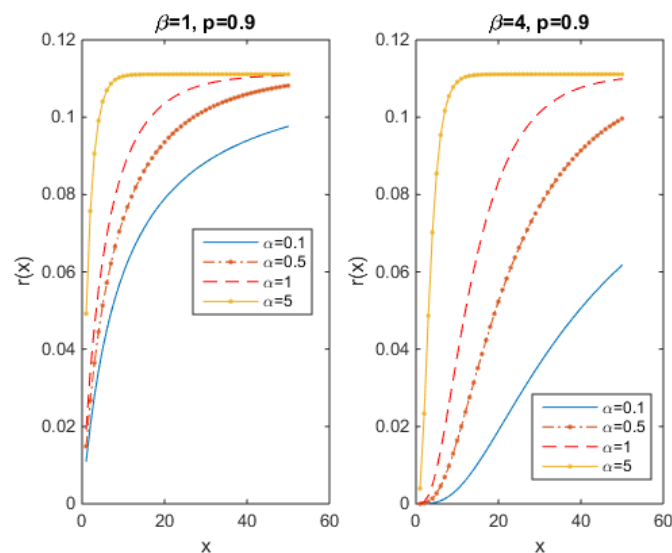
for $x = 1, 2, ...$.

Fig. 4. Failure rate function of a $GWG(\alpha, \beta, p)$ distribution for fixed $p = 0.9$ and different values of α and β.
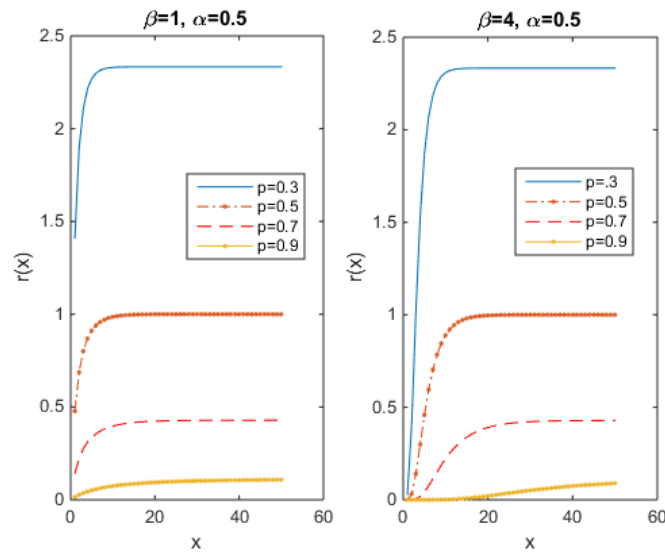
Fig. 5. Failure rate function of a $GWG(\alpha, \beta, p)$ distribution for fixed $\alpha = 0.5$ and different values of β and p.

Remark 3.2. Since the failure rate function is non-decreasing, the mean residual life function $m(x)$ is non-increasing of x.

Figures 6 and 7 illustrate the mean residual functions of a GWG distribution for different values of α, β and p.
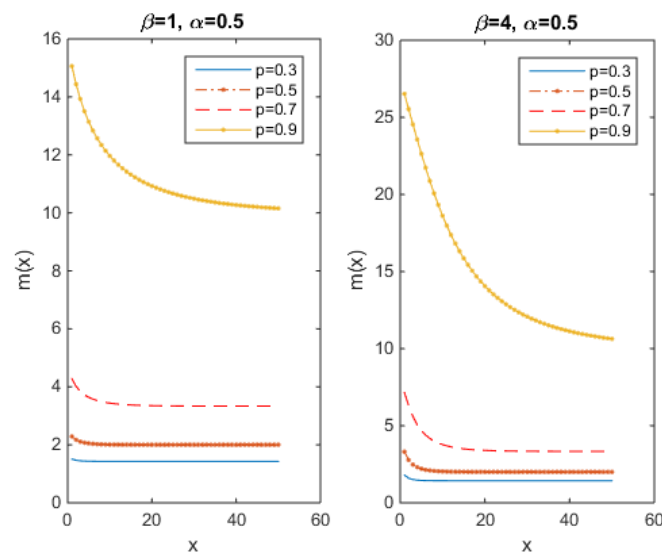

Fig. 6. Mean residual function of $GWG(\alpha, \beta, p)$ distribution for fixed $\alpha = 0.5$ and different values of β and p.
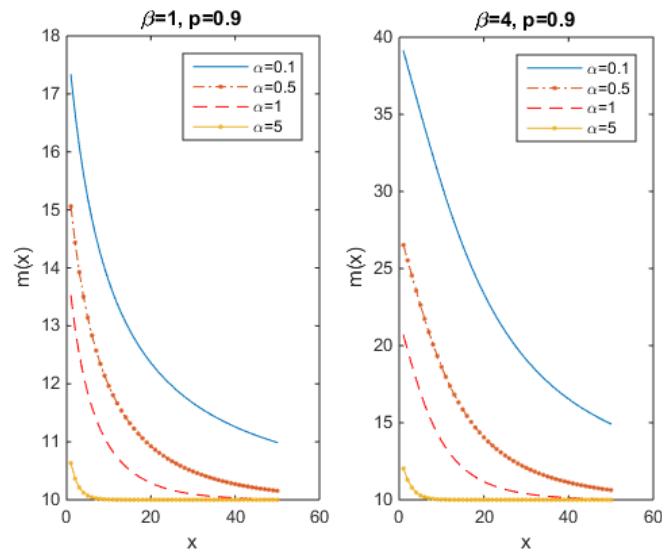
Fig. 7. Mean residual function of $GWG(\alpha, \beta, p)$ distribution for fixed $p = 0.9$ and different values of α and β.

3.4. *Stochastic orders*

Stochastic ordering is an important measure to judge comparative behaviours of random variables. There are many types of stochastic orderings with different implications and applications (see Shaked and Shanthikumar (2007) for detail). In this note, first we give a short review of definitions of the required stochastic orders: the likelihood ratio order ($\leq_{lr}$), the usual stochastic order ($\leq_{st}$), the hazard rate order ($\leq_{hr}$) and the expectation order ($\leq_E$).

Definition 3.1. Let X and Y be two discrete random variables with probability mass functions p_X and p_Y, respectively. Then, we have

(a) $X \leq_{lr} Y$ if $p_Y(x)/p_X(x)$ is an increasing function in x.
(b) $X \leq_{st} Y$ if $P(X \geq x) \leq P(Y \geq x)$ for all x.
(c) $X \leq_{hr} Y$ if $p_Y(x)/P(Y \geq x) \leq p_X(x)/P(X \geq x)$ for all x.
(d) $X \leq_E Y$ if $E(X) \leq E(Y)$.

Theorem 3.4. *Let* $X \sim G_1(1-p)$ *and* $Y \sim GWG(\alpha, \beta, p)$. *Then,* $X \leq_{lr} (\leq_{hr}, \leq_{st}$ *and* $\leq_E)Y$.

Proof. Since

$$\frac{p_Y(x)}{p_X(x)} = \frac{A^{-1}(1-p^{\alpha x})^{\beta}}{(1-p)},$$

is an increasing function in x, then $X \leq_{lr} Y$. Now, based on the well-known results of Shaked and Shanthikumar (2007), it follows that $X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y \Rightarrow X \leq_E Y$. As required. $\square$

3.5. *Order statistics*

Order statistics are among the most fundamental tools in non-parametric statistics and inference. Here, we discuss some properties of order statistics for the discrete GWG distribution. Suppose that

$X_1, ..., X_m$ is a random sample from $GWG(\alpha, \beta, p)$ and $X_{1:m}, ..., X_{m:m}$ are the corresponding order statistics. Then, the pmf and the cdf of $X_{i:m}$ for any $x \in N$ can be expressed as

$$P(X_{i:m} = x) = \frac{m!}{(i-1)!(m-i)!} \int_{F(x-1)}^{F(x)} u^{i-1}(1-u)^{(m-i)} du,$$

$$= \frac{m!}{(i-1)!(m-i)!} \sum_{k=0}^{m-i} (-1)^k \binom{m-i}{k} \int_{F(x-1)}^{F(x)} u^{i+k-1} du,$$

$$= \frac{m!}{(i-1)!(m-i)!} \sum_{k=0}^{m-i} \binom{m-i}{k} \frac{(-1)^k}{(i+k)} \left\{ \sum_{t=0}^{i+k} A^{-t} \binom{i+k}{t} (-1)^t p^{[x]t} \left\{ \sum_{j=0}^{\infty} A_j p^{\alpha[x]j} \right\}^t, \right.$$

$$\left. - \sum_{t=0}^{i+k} A^{-t} \binom{i+k}{t} (-1)^t p^{[x-1]t} \left\{ \sum_{j=0}^{\infty} A_j p^{\alpha[x-1]j} \right\}^t \right\},$$

$$= \frac{m!}{(i-1)!(m-i)!} \sum_{k=0}^{m-i} \binom{m-i}{k} \frac{(-1)^k}{(i+k)} \left\{ \sum_{j=0}^{\infty} \sum_{t=0}^{i+k} d_{t,j} A^{-t} \binom{i+k}{t} (-1)^t p^{[x]t + \alpha[x]j}, \right.$$

$$\left. - \sum_{j=0}^{\infty} \sum_{t=0}^{i+k} d_{t,j} A^{-t} \binom{i+k}{t} (-1)^t p^{[x-1]t + \alpha[x-1]j} \right\},$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{m-i} \sum_{t=0}^{i+k} \frac{d_{t,j} A^{-t} (-1)^{t+k} m! (i+k-1)!}{(i-1)! t! (i+k-t)! k! (m-i-k)!} \left\{ p^{(t+\alpha j)[x]} - p^{(t+\alpha j)[x-1]} \right\} \quad (3.11)$$

and

$$P(X_{i:m} \le x) = \sum_{l=i}^{m} \binom{m}{l} (F(x))^l (1 - F(x))^{m-l},$$

$$= \frac{m!}{(i-1)!(m-i)!} \int_0^{F(x)} u^{i-1}(1-u)^{(m-i)} du,$$

$$= \frac{m!}{(i-1)!(m-i)!} \sum_{k=0}^{m-i} \binom{m-i}{k} \frac{(-1)^k}{(i+k)} \left[1 - A^{-1} \sum_{j=0}^{n} A_j p^{(\alpha j+1)[x]} \right]^{i+k},$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{m-i} \sum_{t=0}^{i+k} \frac{d_{t,j} A^{-t} (-1)^{t+k} m! (i+k-1)!}{(i-1)! t! (i+k-t)! k! (m-i-k)!} p^{(t+\alpha j)[x]} \quad (3.12)$$

respectively.

Moments of order statistics play an important role in quality control testing and reliability. Now, we provide the rth-order moment of $X_{i:m}$ from a GWG distribution as follows

$$E(X_{i:m}^r) = \sum_{x=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{m-i} \sum_{t=0}^{i+k} x^r \frac{d_{t,j} A^{-t} (-1)^{t+k} m! (i+k-1)!}{(i-1)! t! (i+k-t)! k! (m-i-k)!} \left\{ p^{(t+\alpha j)[x]} - p^{(t+\alpha j)[x-1]} \right\},$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{m-i} \sum_{t=0}^{i+k} \frac{d_{t,j} A^{-t} (-1)^{t+k} m! (i+k-1)!}{(i-1)! t! (i+k-t)! k! (m-i-k)!} \left\{ \sum_{x=1}^{\infty} x^r p^{(t+\alpha j)x} - \sum_{x=1}^{\infty} x^r p^{(t+\alpha j)(x-1)} \right\},$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{m-i} \sum_{t=0}^{i+k} \sum_{x=1}^{r} \frac{d_{t,j} A^{-t} (-1)^{t+k} m! (i+k-1)! S(r,x) x!}{(i-1)! t! (i+k-t)! k! (m-i-k)!} \times \frac{p^{(t+\alpha j)(x-1)} (p^{t+\alpha j} - 1)}{(1 - p^{t+\alpha j})^{(x+1)}}, \quad (3.13)$$

where $S(r,x)$ can be obtained from Eq. (3.3).

3.6. *Entropies*

An entropy is a measure of unpredictability or uncertainty of a random variable. It has been used in various situations in science and engineering. Three of the most popular entropies and information indices are the Renyi, Mathai-Haubold and Shannon entropies. For a discrete random variable X with pmf $p(.)$, Renyi, Mathai-Haubold and Shannon entropies are defined as

$$J_R(\gamma) = \frac{1}{1-\gamma} \log\left\{ \sum_{x=1}^{\infty} p(x)^{\gamma} \right\},$$

$$J_{MH}(\delta) = \frac{\sum_{x=1}^{\infty} \left[p(x) \right]^{2-\delta} - 1}{\delta - 1}$$

and $E\left\{ -\log p(X) \right\}$, respectively, where $\gamma > 0$, $\gamma \neq 1$, $\delta \neq 1$ and $\delta < 2$. The Mathai-Haubold entropy is an inaccuracy measure through disturbance or distortion of a systems. In what follows, we obtain their explicit forms for a GWG distribution.

Theorem 3.5. *The Renyi entropy of the $GWG(\alpha, \beta, p)$ distribution is*

$$J_R(\gamma) = \frac{1}{1-\gamma} \log\left[\sum_{k=0}^{\infty} \binom{\gamma\beta}{j} (-1)^k \frac{A^{-\gamma} p^{\alpha k}}{1 - p^{(\alpha k + \gamma)}} \right]. \tag{3.14}$$

Proof. Using Taylor series expansion

$$(1-x)^l = \sum_{j=0}^{\infty} (-1)^j \binom{l}{j} x^j,$$

for $l \in \mathscr{R}$ and $|x| < 1$ we have

$$\sum_{x=1}^{\infty} p(x)^{\gamma} = \sum_{x=1}^{\infty} A^{-\gamma} p^{\gamma x - \gamma} (1 - p^{\alpha x})^{\gamma\beta},$$

$$= \sum_{k=0}^{\infty} \sum_{x=1}^{\infty} \binom{\gamma\beta}{k} (-1)^k A^{-\gamma} p^{(\alpha k + \gamma)x - \gamma},$$

$$= \sum_{k=0}^{\infty} \binom{\gamma\beta}{j} (-1)^k \frac{A^{-\gamma} p^{\alpha k}}{1 - p^{(\alpha k + \gamma)}}. \tag{3.15}$$

Thus, the proof is straightforward. □

Theorem 3.6. *The Mathai-Haubold entropy of a $GWG(\alpha, \beta, p)$ distribution is*

$$J_{MH}(\delta) = \frac{1}{\delta - 1} \left[\sum_{k=0}^{\infty} \binom{2\beta - \delta\beta}{k} (-1)^k \frac{A^{\delta-2} p^{(\alpha k)}}{1 - p^{(\alpha k + 2 - \delta)}} - 1 \right]. \tag{3.16}$$

Proof. Since

$$\sum_{x=1}^{\infty} p(x)^{2-\delta} = \sum_{k=0}^{\infty} \binom{2\beta - \delta\beta}{k} (-1)^k \frac{A^{\delta-2} p^{(\alpha k)}}{1 - p^{(\alpha k + 2 - \delta)}},$$

the proof is complete. □

Theorem 3.7. *The Shannon entropy of a $GWG(\alpha,\beta,p)$ distribution is*

$$E\{-\log p(X)\} = \log A + (A^{-1}\sum_{j=0}^{\infty}(-1)^j\binom{\beta}{j}\frac{p^{\alpha j}}{(1-p^{\alpha j+1})^2} - 1)(\log p)$$

$$- \beta A^{-1}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{p^{\alpha(k+j+1)}}{(k+1)(1-p^{\alpha(k+j+1)+1})}. \tag{3.17}$$

Proof.

$$E\{-\log p(X)\} = \log A + (E(X)-1)(\log p) + \beta E(\log(1-p^{\alpha X})),$$

$$= \log A + (E(X)-1)(\log p) - \beta\sum_{k=0}^{\infty}\frac{E(p^{\alpha(k+1)X})}{k+1},$$

$$= \log A + (A^{-1}\sum_{j=0}^{\infty}(-1)^j\binom{\beta}{j}\frac{p^{\alpha j}}{(1-p^{\alpha j+1})^2} - 1)(\log p)$$

$$- \beta A^{-1}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{p^{\alpha(k+j+1)}}{(k+1)(1-p^{\alpha(k+j+1)+1})}. \tag{3.18}$$

$\square$

Remark 3.3. We may observe that Shannon entropy (3.17) is a particular case of the Renyi entropy (3.14) when $\gamma\uparrow 1$.

3.7. *Stress-strength reliability function$(R = P(X < Y))$*

In the context of reliability, $R = P(X < Y)$ is widely known as the stress-strength of a model: if X denotes the stress that a system is subjected to and Y is the strength of the system, then $R = P(X < Y)$ is the probability of the reliability of the system. Now, suppose that $X \sim GWG(\alpha_1,\beta_1,p_1)$ and $Y \sim GWG(\alpha_2,\beta_2,p_2)$ are two independent random variables. Then, the stress-strength reliability function $R = P(X < Y)$ is given by

$$R = P(X < Y) = \sum_{x=1}^{\infty}P(Y > x)P(X = x),$$

$$= \sum_{x=1}^{\infty}A_2^{-1}\sum_{j=0}^{\infty}A_{2j}p_2^{(\alpha_2 j+1)x}A_1^{-1}p_1^{x-1}(1-p_1^{\alpha_1 x})^{\beta_1},$$

$$= \frac{(A_1 A_2)^{-1}}{p_1}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}A_{2j}(-1)^l\binom{\beta_1}{l}\sum_{x=1}^{\infty}(p_2^{\alpha_2 j+1}p_1^{\alpha_1 l+1})^x,$$

$$= (A_1 A_2)^{-1}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}A_{2j}(-1)^l\binom{\beta_1}{l}\frac{p_2^{\alpha_2 j+1}p_1^{\alpha_1 l}}{1-p_2^{\alpha_2 j+1}p_1^{\alpha_1 l+1}}, \tag{3.19}$$

where $A_{2j} = (-1)^j\binom{\beta_2}{j}\frac{p_2^{\alpha_2 j}}{1-p_2^{\alpha_2 j+1}}$.

The stress-strength reliability function (3.19) for different values of α_1, β_1, p_1 and α_2, β_2, p_2 are computed in Table 1 below for illustration. As we see, R is increasing as α_1, β_2 and p_2 increase and it is decreasing as α_2, β_1 and p_1 increase.

Table 1. $R = P(X < Y)$ for different values of (α_1, β_1, p_1) and (α_2, β_2, p_2)

β_1				1				
β_2			1				4	
$(\alpha_2, p_2)/(\alpha_1, p_1)$	(1, 0.3)	(1, 0.9)	(4, 0.3)	(4, 0.9)	(1, 0.3)	(1, 0.9)	(4, 0.3)	(4, 0.9)
(1, 0.3)	0.2628	0.0128	0.2801	0.0253	0.4238	0.0219	0.4500	0.0430
(1, 0.9)	0.9620	0.4824	0.9655	0.5754	.9992	0.7011	0.9993	0.7808
(4, 0.3)	0.2171	0.0104	0.2317	0.0206	0.2208	0.0106	0.2356	0.0209
(4, 0.9)	0.9255	0.3867	0.9317	0.4779	0.9862	0.4736	0.9884	0.5745
β_1				4				
β_2			1				4	
$(\alpha_2, p_2)/(\alpha_1, p_1)$	(1, 0.3)	(1, 0.9)	(4, 0.3)	(4, 0.9)	(1, 0.3)	(1, 0.9)	(4, 0.3)	(4, 0.9)
(1, 0.3)	0.2018	0.0002	0.2787	0.0044	0.3315	0.0004	0.4479	0.0080
(1, 0.9)	0.9489	0.2738	0.9653	0.4889	0.9986	0.4854	0.9993	0.7216
(4, 0.3)	0.1657	0.0001	0.2305	0.0035	0.1685	0.0002	0.2344	0.0035
(4, 0.9)	0.9028	0.1977	0.9312	0.3859	0.9775	0.2522	0.9883	0.4793

4. WG distribution as a special case

In this section, we consider the weighted geometric (WG) distribution which was proposed by Bhati and Joshi (2015) as a special case of GWG distribution when $\beta = 1$ and present some more characterizations and convolution properties for the distribution in such a case. Thus, a random variable X has a WG distribution if its pmf takes the form

$$P(X = x) = \frac{(1-p)(1-p^{\alpha+1})}{(1-p^{\alpha})}p^{x-1}(1-p^{\alpha x}), \quad x \in N = \{1, 2, ...\}. \tag{4.1}$$

where $\alpha > 0$ and $0 < p < 1$. Infact, WG distribution is a discrete version of WE distribution (1.3), which was constructed by inserting the weighted exponential density (1.3) into (1.1).

Remark 4.1. We can easily observe that

(i) The WG distribution reduces to

$$p_1(x) = P(X = x) = (1-p)p^{x-1}, \quad x = 1, 2, ..., \tag{4.2}$$

i.e., the classic geometric distribution $G_1(1-p)$, when $\alpha \longrightarrow \infty$.

(ii) When $\alpha \longrightarrow 0$, WG distribution reduces to

$$p_1(x) = P(X = x) = \frac{(1-p)^2}{p}xp^x, \quad x = 1, 2, ..., \tag{4.3}$$

which is the pmf of the discrete gamma distribution with parameters 2 and p.

The corresponding cdf and failure rate of the WG distribution are

$$F(x) = 1 - \frac{\left\{1 - p^{\alpha+1} - p^{\alpha([x]+1)}(1-p)\right\}}{1 - p^{\alpha}}p^{[x]} \quad x \geq 1, \tag{4.4}$$

and

$$r(x) = \frac{(1-p)(1-p^{\alpha+1})p^{x-1}(1-p^{\alpha x})}{\left\{1 - p^{\alpha+1} - p^{\alpha(x+1)}(1-p)\right\}p^x}, \quad x = 1, 2, ..., \tag{4.5}$$

respectively.

Remark 4.2. We may note that the failure rate function $r(x)$ of a $WG(\alpha, p)$ is bounded and we have

$$0 < r(x) < \frac{1-p}{p},$$

where $\frac{1-p}{p}$ is the failure rate of the $G_1(1-p)$ distribution.

Also, the moment generating function and mean residual life time of a $WG(\alpha, p)$ distribution are

$$M_X(t) = \frac{(1-p)(1-p^{\alpha+1})e^t}{(1-pe^t)(1-p^{\alpha+1}e^t)}. \tag{4.6}$$

and

$$m(x) = E(X - x|X > x) = \frac{(1-p^{\alpha+1})^2 - (1-p)^2 p^{\alpha(x+1)}}{\{1 - p^{\alpha+1} - p^{\alpha(x+1)}(1-p)\}(1-p^{\alpha+1})(1-p)}, \tag{4.7}$$

respectively.

Remark 4.3. It is interesting to observe that the mean residual life function is also bounded and we have

$$\frac{(1-p^{\alpha+2})}{(1-p)(1-p^{\alpha+1})} < m(x) < \frac{1}{1-p}.$$

Bhati and Joshi (2015) by using mgf (4.6) for a WG distribution indicated that a discrete random variable X is $WG(\alpha, p)$ if, and only if, it can be decomposed as

$$X \overset{d}{=} U + V \tag{4.8}$$

where $U \sim G_0(1-p)$ and $V \sim G_1(1-p^{\alpha+1})$ are two independent random variables. Based on (4.8), since discrete strongly unimodal distributions are closed under convolutions and geometric distribution is strongly unimodal, then, $WG(\alpha, p)$ distribution is strongly unimodal and thus it is unimodal which was obviously expected by Theorem 3.1.

Now, let $G_0(1-p)$, $nb_0(r, p)$ and $nb_1(r, p)$ represent geometric, negative binomial of the first type and negative binomial of the second type distributions with probability mass functions

$$p_0(x) = (1-p)p^x, \qquad x = 0, 1, 2, ..., \tag{4.9}$$

$$p_{0,r}(x) = \binom{x+r-1}{x} p^x (1-p)^r, \qquad x = 0, 1, 2, ..., \tag{4.10}$$

and

$$p_{1,r}(x) = \binom{x-1}{x-r} p^{x-r} (1-p)^r, \qquad x = r, r+1, r+2, ..., \tag{4.11}$$

respectively. Then, by considering Eq.'s (3.1) and (4.8) the following convolution properties are immediate.

Theorem 4.1. *Suppose that $Y_1 \sim nb_0(n, 1-p)$ and $Y_2 \sim nb_1(n, 1-p^{\alpha+1})$ are two independent random variables and $X_1, X_2, ..., X_n$ is a random sample of size n from X. Then for $\alpha > 0$*

$$X_1 + X_2 + ... + X_n \overset{d}{=} Y_1 + Y_2 \tag{4.12}$$

if, and only if, $X \sim WG(\alpha, p)$.

Proof. The proof is straightforward. □

Theorem 4.2. *Suppose that $Y_1 \sim G_0(1-p)$, $Y_2 \sim G_0(1-p^{\alpha+1})$, $Y_3 \sim G_1(1-p^{2\alpha+1})$ and $Y_4' \sim Ber(\frac{p^{\alpha+1}}{1+p^{\alpha+1}})$ are independent random variables. Then, for $\alpha > 0$*

$$X \overset{d}{=} Y_1 + Y_2 + Y_3 + Y_4' \tag{4.13}$$

if, and only if, $X \sim GWG(\alpha, 2, p)$.

Proof. By mgf (3.1), for $\beta = 2$ we have

$$
\begin{aligned}
M_X(t) &= \frac{(1-p)(1-p^{\alpha+1})(1-p^{2\alpha+1})(1+p^{\alpha+1}e^t)e^t}{(1-pe^t)(1-p^{\alpha+1}e^t)(1-p^{2\alpha+1}e^t)(1+p^{\alpha+1})} \\
&= M_{Y_1}(t)M_{Y_2}(t)M_{Y_3}(t)M_{Y_4'}(t)
\end{aligned}
\tag{4.14}
$$

which implies the result. □

Note that Eq.'s (4.8) and (4.13) can be used for generating $GWG(\alpha, 1, p)$ and $GWG(\alpha, 2, p)$ distributions, respectively.

5. Maximum Likelihood Estimation(MLE)

In this section, we discuss estimation of parameters of a $GWG(\alpha, \beta, p)$ distribution, by the maximum likelihood method. Let $x_1, ..., x_m$ be m observations of a random sample from a GWG distribution with the vector parameter $\Theta = (\alpha, \beta, p)$. If we rewrite pmf of a $GWG(\alpha, \beta, p)$ distribution as

$$p(x) = P(X = x) = \frac{p^x(1-p^{\alpha x})^\beta}{\sum_{t=1}^{\infty} p^t(1-p^{\alpha t})^\beta},$$

the log-likelihood function for $\Theta = (\alpha, \beta, p)$ is given by

$$\ell_m(\Theta) = \sum_{i=1}^{m} x_i \log p + \beta \sum_{i=1}^{m} \log(1-p^{\alpha x_i}) - m \log \sum_{t=1}^{\infty} p^t(1-p^{\alpha t})^\beta.$$

Therefore the first derivatives of $\ell_m(\Theta)$ with respect to the parameters α, β and p are:

$$\frac{\partial \ell_m(\Theta)}{\partial \alpha} = -\beta \sum_{i=1}^{m} \frac{x_i p^{\alpha x_i} \log p}{1-p^{\alpha x_i}} + m \frac{\sum_{t=1}^{\infty} t\beta p^{t(\alpha+1)}(1-p^{\alpha t})^{\beta-1} \log p}{\sum_{t=1}^{\infty} p^t(1-p^{\alpha t})^\beta} = 0,$$

$$\frac{\partial \ell_m(\Theta)}{\partial \beta} = \sum_{i=1}^{m} \log(1-p^{\alpha x_i}) - m \frac{\sum_{t=1}^{\infty} p^t(1-p^{\alpha t})^\beta \log(1-p^{\alpha t})}{\sum_{t=1}^{\infty} p^t(1-p^{\alpha t})^\beta} = 0,$$

and

$$\frac{\partial \ell_m(\Theta)}{\partial p} = \frac{\sum_{i=1}^{m} x_i}{p} - \beta \sum_{i=1}^{m} \frac{\alpha x_i p^{\alpha x_i - 1}}{1 - p^{\alpha x_i}} - m \frac{\sum_{t=1}^{\infty} t p^{t-1}(1 - p^{\alpha t})^{\beta-1}(1 - p^{\alpha t} - \beta \alpha p^{\alpha t})}{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}} = 0.$$

Solutions of above equations yeild the MLE of $\Theta = (\alpha, \beta, p)$ which can be obtained using the function "optim" in the statistical software R. For interval estimation and hypothesis testings of the model parameters, we require the information matrix $I(\theta) = E[-\frac{\partial^2 \ell_n(\Theta)}{\partial \Theta^2}]$, where

$$\frac{\partial^2 \ell_n(\Theta)}{\partial \Theta^2} = \begin{bmatrix} \frac{\partial^2 \ell_n(\Theta)}{\partial \alpha^2} & \frac{\partial^2 \ell_n(\Theta)}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell_n(\Theta)}{\partial \alpha \partial p} \\ \frac{\partial^2 \ell_n(\Theta)}{\partial \beta \partial \alpha} & \frac{\partial^2 \ell_n(\Theta)}{\partial \beta^2} & \frac{\partial^2 \ell_n(\Theta)}{\partial \beta \partial p} \\ \frac{\partial^2 \ell_n(\Theta)}{\partial p \partial \alpha} & \frac{\partial^2 \ell_n(\Theta)}{\partial p \partial \beta} & \frac{\partial^2 \ell_n(\Theta)}{\partial p^2} \end{bmatrix},$$

with

$$\frac{\partial^2 \ell_n(\Theta)}{\partial \alpha^2} = -\beta \sum_{i=1}^{m} \frac{x_i^2 p^{\alpha x_i}(\log p)^2}{(1 - p^{\alpha x_i})^2} + m \frac{\sum_{t=1}^{\infty} t^2 \beta p^{t(\alpha+1)}(1 - p^{\alpha t})^{\beta-2}(\log p)^2(1 - \beta p^{\alpha t})}{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}},$$
$$+ m \frac{[\sum_{t=1}^{\infty} t \beta p^{t(\alpha+1)}(1 - p^{\alpha t})^{\beta-1} \log p]^2}{[\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}]^2},$$

$$\frac{\partial^2 \ell_n(\Theta)}{\partial \beta^2} = -m \frac{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}[\log(1 - p^{\alpha t})]^2}{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}} + m \frac{[\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta} \log(1 - p^{\alpha t})]^2}{[\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}]^2},$$

$$\frac{\partial^2 \ell_n(\Theta)}{\partial p^2} = -\frac{\sum_{i=1}^{m} x_i}{p^2} - \beta \sum_{i=1}^{m} \frac{\alpha x_i(\alpha x_i - 1) p^{\alpha x_i - 2}}{1 - p^{\alpha x_i}} - \beta \sum_{i=1}^{m} \frac{(\alpha x_i)^2 p^{2\alpha x_i - 2}}{(1 - p^{\alpha x_i})^2}$$
$$- m \frac{\sum_{t=1}^{\infty} (1 - \beta) \alpha t^2 p^{\alpha t + t - 2}(1 - p^{\alpha t})^{\beta-2}(1 - p^{\alpha t} - \beta \alpha p^{\alpha t})}{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}},$$
$$- m \frac{\sum_{t=1}^{\infty} [(t-1) p^{t-2} - (\alpha t + t - 1) p^{\alpha t + t - 2} - \beta \alpha(\alpha t + t - 1) p^{\alpha t + t - 2}][t(1 - p^{\alpha t})^{\beta-1}]}{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}},$$
$$+ m \frac{[\sum_{t=1}^{\infty} t p^{t-1}(1 - p^{\alpha t})^{\beta-1}(1 - p^{\alpha t} - \alpha \beta p^{\alpha t})]^2}{[\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}]^2},,$$

Table 2. MLEs and standard errors of α, β and p in a $GWG(\alpha, \beta, p)$ distribution for different values of m and $\beta = 1$.

α	.5							
p	0.3				0.9			
m	100	500	1000	5000	100	500	1000	5000
$\hat{\alpha}$	0.6623	0.3920	0.4578	0.4845	0.6005	0.5153	0.4576	0.4843
$\hat{\beta}$	1.3680	1.2967	1.1006	1.0615	1.1357	1.0873	1.0431	1.0086
$\hat{p}$	0.2954	0.2772	0.2773	0.2940	0.8951	0.8953	0.8968	0.8985
$\widehat{S.E}(\hat{\alpha})$	0.1618	0.1226	0.1002	0.0315	0.3994	0.1365	0.0814	0.0118
$\widehat{S.E}(\hat{\beta})$	0.0931	0.0301	0.0330	0.0051	0.1045	0.0284	0.0092	0.0012
$\widehat{S.E}(\hat{p})$	0.0467	0.0411	0.0320	0.0131	0.0071	0.0079	0.0056	0.0003
α	1							
p	0.3				0.9			
m	100	500	1000	5000	100	500	1000	5000
$\hat{\alpha}$	1.6840	0.9267	0.9453	1.0146	1.1607	1.2100	1.0129	0.9526
$\hat{\beta}$	1.9557	1.2489	1.1082	1.0212	1.2956	1.0790	1.0887	0.9980
$\hat{p}$	0.2983	0.2948	0.2875	0.2992	0.8900	0.9023	0.8965	0.8989
$\widehat{S.E}(\hat{\alpha})$	0.3487	0.0614	0.0328	0.0180	0.5781	0.1139	0.0730	0.0154
$\widehat{S.E}(\hat{\beta})$	0.1759	0.0307	0.0126	0.0053	0.1916	0.0249	0.0152	0.0022
$\widehat{S.E}(\hat{p})$	0.0052	0.0043	0.0022	0.0009	0.0102	0.0015	0.0018	0.0002
α	5							
p	0.3				0.9			
m	100	500	1000	5000	100	500	1000	5000
$\hat{\alpha}$	7.5890	6.3035	5.6611	5.2121	4.9608	5.5291	5.2609	5.0578
$\hat{\beta}$	2.0001	1.3491	1.2546	1.0040	1.2874	1.1145	1.0502	1.0066
$\hat{p}$	0.2758	0.2914	0.2972	0.2977	0.8963	0.8990	0.9003	0.9002
$\widehat{S.E}(\hat{\alpha})$	0.9889	0.6733	0.2890	0.1842	0.2979	0.2372	0.1911	0.0308
$\widehat{S.E}(\hat{\beta})$	0.7762	0.5082	0.4806	0.0313	0.0776	0.0516	0.0208	0.0040
$\widehat{S.E}(\hat{p})$	0.0307	0.0029	0.0006	0.0002	0.0012	0.0009	0.0005	0.0000

$$\frac{\partial^2 \ell_n(\Theta)}{\partial \alpha \partial \beta} = -\sum_{i=1}^{m} \frac{x_i p^{\alpha x_i} \log p}{1 - p^{\alpha x_i}} + m \frac{\sum_{t=1}^{\infty} t p^{t(\alpha+1)} \log p \{((1-p^{\alpha t})^{\beta-1}) + \beta(1-p^{\alpha t})^{\beta-1} \log(1-p^{\alpha t})\}}{\sum_{t=1}^{\infty} p^t (1-p^{\alpha t})^{\beta}},$$

$$- m \frac{[\sum_{t=1}^{\infty} p^t (1-p^{\alpha t})^{\beta} \log(1-p^{\alpha t})] \times [\sum_{t=1}^{\infty} t \beta p^{t(\alpha+1)} (1-p^{\alpha t})^{\beta-1} \log p]}{(\sum_{t=1}^{\infty} p^t (1-p^{\alpha t})^{\beta})^2},$$

$$\frac{\partial^2 \ell_n(\Theta)}{\partial p \partial \alpha} = m \frac{\sum_{t=1}^{\infty} t^2 p^{(\alpha+1)t-1} \log p (\beta-1)(1-p^{\alpha t})^{\beta-2}(1-p^{\alpha t} - \beta \alpha p^{\alpha t})}{\sum_{t=1}^{\infty} p^t (1-p^{\alpha t})^{\beta}},$$

$$+ m \frac{\sum_{t=1}^{\infty} t p^{t-1} \{t p^{\alpha t} \log p + \beta p^{\alpha t} + \beta \alpha t p^{\alpha t} \log p\} \{(1-p^{\alpha t})^{\beta-1}\}}{\sum_{t=1}^{\infty} p^t (1-p^{\alpha t})^{\beta}},$$

$$+ m \frac{\{\sum_{t=1}^{\infty} t p^{(\alpha+1)t} \beta \log p (1-p^{\alpha t})^{\beta-1}\}\{\sum_{t=1}^{\infty} t p^{t-1} (1-p^{\alpha t})^{\beta-1}(1-p^{\alpha t} - \beta \alpha p^{\alpha t})\}}{(\sum_{t=1}^{\infty} p^t (1-p^{\alpha t})^{\beta})^2},$$

$$- \beta \sum_{i=1}^{m} \frac{x_i p^{\alpha x_i - 1}(1 + x_i \alpha \log p - p^{\alpha x_i})}{(1-p^{\alpha x_i})^2}$$

and

Table 3. MLEs and standard errors of α, β and p in a $GWG(\alpha, \beta, p)$ distribution for different values of m and $\beta = 2$.

α				0.5				
p		0.3					0.9	
m	100	500	1000	5000	100	500	1000	5000
$\hat{\alpha}$	0.8608	0.6421	0.4843	0.5265	0.7211	0.4262	0.4557	0.5248
$\hat{\beta}$	2.0585	2.0612	2.0525	2.0532	2.0708	2.0551	2.0563	2.0524
$\hat{p}$	0.3156	0.3023	0.2890	0.3018	0.8992	0.8926	0.8914	0.9009
$\widehat{S.E}(\hat{\alpha})$	0.3908	0.0967	0.0327	0.0084	0.2851	0.0612	0.0432	0.0060
$\widehat{S.E}(\hat{\beta})$	0.2168	0.1162	0.0529	0.0238	0.4453	0.0796	0.0599	0.0236
$\widehat{S.E}(\hat{p})$	0.0348	0.0130	0.0057	0.0013	0.0110	0.0033	0.0026	0.0002
α				1				
p		0.3					0.9	
m	100	500	1000	5000	100	500	1000	5000
$\hat{\alpha}$	1.6928	1.0801	0.9367	0.9803	0.7833	0.9012	1.0248	1.0376
$\hat{\beta}$	2.1716	2.0576	2.0499	2.0546	2.0585	2.0539	2.0561	2.0530
$\hat{p}$	0.2698	0.3050	0.2931	0.2977	0.8806	0.8935	0.9000	0.9010
$\widehat{S.E}(\hat{\alpha})$	0.2486	0.0691	0.0415	0.0094	0.3962	0.0743	0.0319	0.0058
$\widehat{S.E}(\hat{\beta})$	0.6727	0.0914	0.0558	0.0247	0.2160	.0763	.0592	0.0237
$\widehat{S.E}(\hat{p})$	0.0472	0.0063	0.0045	0.0008	0.0159	0.0026	0.0008	0.0001
α				5				
p		0.3					0.9	
m	100	500	1000	5000	100	500	1000	5000
$\hat{\alpha}$	6.8484	6.5166	6.3905	5.8762	4.2699	5.3664	4.8635	5.0921
$\hat{\beta}$	2.2124	2.1181	2.0149	1.9823	2.0636	2.0462	2.0430	2.0510
$\hat{p}$	0.2724	0.2855	0.2861	0.2948	0.8911	0.8987	0.8991	0.9004
$\widehat{S.E}(\hat{\alpha})$	0.3551	0.2210	0.1549	0.0181	0.8742	0.2110	0.1014	0.0194
$\widehat{S.E}(\hat{\beta})$	0.4979	0.1821	0.1326	0.0358	0.3013	0.0969	0.0846	0.0240
$\widehat{S.E}(\hat{p})$	0.0125	0.0043	0.0038	0.0003	0.0081	0.0010	0.0004	0.0001

$$\frac{\partial^2 \ell_n(\Theta)}{\partial p \partial \beta} = -\sum_{i=1}^{m} \frac{\alpha x_i p^{\alpha x_i - 1}}{1 - p^{\alpha x_i}} - m \frac{\sum_{t=1}^{\infty} t p^{t-1} (1 - p^{\alpha t})^{\beta - 1} \{(1 - p^{\alpha t} - \alpha \beta p^{\alpha t}) \log(1 - p^{\alpha t}) - \alpha p^{\alpha t}\}}{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta}},$$

$$- m \frac{\{\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta} \log(1 - p^{\alpha t})\}\{\sum_{t=1}^{\infty} t p^{t-1} (1 - p^{\alpha t})^{\beta - 1} (1 - p^{\alpha t} - \beta \alpha p^{\alpha t})\}}{(\sum_{t=1}^{\infty} p^t (1 - p^{\alpha t})^{\beta})^2}.$$

Let MLE of $\Theta = (\alpha, \beta, p)$ be $\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{p})$, then for large m, by assuming the regularity conditions, the distribution of $(\hat{\alpha} - \alpha, \hat{\beta} - \beta, \hat{p} - p)$ can be approximated by a bivariate normal distribution with the mean vector zero and variance-covariance matrix I^{-1}.

Tables 2 and 3 show the MLEs and standard errors of parameters of $DWE(\alpha, \beta, p)$ distribution for different values of m and $\beta = 1$ and $\beta = 2$, respectively. We generate these distributions by using Eq.'s (4.8) and (4.13), respectively. As we see, from these tables, as m increases, we get better estimates for α, β and p.

6. Application

In this section, we apply a real data set (cf. Makčutek (2008)), concerning rank frequencies of graphemes in a Slavic language: Slovene, indicated in Table 4, to investigate the advantage of our proposed model. As mentioned previously, Nekoukhou et al. (2011) and Nekoukhou et al. (2015)

Table 4. Observed and expected frequencies of $GWG(\alpha,\beta,p), DBE(\gamma,\lambda,p), WG(\alpha,p), DGE(\gamma,p)$ and $G_1(1-p)$ models

i	$f(i)$	$\hat{e}_{GWG}(i)$	$\hat{e}_{DBE}(i)$	$\hat{e}_{WG}(i)$	$\hat{e}_{DGE}(i)$	$\hat{e}_{G_1}(i)$
1	32036	28561.85	28912.20	29664.93	21531.44	40314.95
2	31891	31171.01	31520.83	34707.36	24339.73	35134.48
3	31122	30401.11	30645.02	32759.53	24571.81	30619.70
4	27150	28316.37	28438.76	29146.71	23690.92	26685.06
5	22905	25717.69	25733.98	25435.06	22278.32	23256.03
6	16088	22987.84	22922.96	22048.21	20625.58	20267.63
7	16084	20322.77	20202.21	19067.53	18893.86	17663.24
8	15221	17822.48	17668.59	16476.12	17175.91	15393.52
9	14668	15533.91	15364.66	14232.72	15525.16	13415.45
10	14043	13473.64	13302.60	12293.51	13971.32	11691.56
11	13034	11640.81	11477.60	10618.12	12529.31	10189.20
12	10517	10024.74	9875.66	9170.93	11204.78	8879.89
13	10514	8609.54	8478.30	7920.95	9997.45	7738.82
14	10216	7377.00	7265.30	6841.33	8903.47	6744.38
15	9568	6308.29	6216.37	5908.86	7916.81	5877.73
16	7446	5385.02	5312.14	5103.48	7030.25	5122.44
17	6413	4589.85	4534.59	4407.88	6236.05	4464.21
18	5361	3906.81	3867.37	3807.08	5526.35	3890.56
19	5055	3321.39	3295.81	3288.18	4893.48	3390.62
20	4608	2820.64	2806.90	2840.03	4330.10	2954.92
21	2606	2393.05	2389.18	2452.91	3829.31	2575.22
22	2554	2028.47	2032.66	2118.57	3384.69	2244.30
23	2463	1718.05	1728.64	1829.81	2990.38	1955.91
24	1675	1454.05	1469.57	1580.41	2640.99	1704.57
25	497	1229.79	1248.95	1365.01	2331.64	1485.54

N=313735

proposed a discrete analogue of the generalize exponential and a discrete beta-exponential distributions, respectively, with pmfs

$$DGE : f(x;\gamma,p) = kp^{x-1}(1-p^x)^{\gamma-1} \qquad x \in N = 1,2,... \qquad (6.1)$$

and

$$DBE : f(x;\gamma,\lambda,p) = cp^{\lambda(x-1)}(1-p^x)^{\gamma-1} \qquad x \in N = 1,2,..., \qquad (6.2)$$

where $\gamma > 0$, $\lambda > 0$, $0 < p < 1$, $k = \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j p^j}{1-p^{1+j}}$ and $c = \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j p^j}{1-p^{\beta+j}}$ in which $\binom{\gamma-1}{j} = \frac{(\gamma-1)(\gamma-2)...(\gamma-j)}{j!}$. They used the data set presented in Table 4 to reveal the capability of their distribution and indicated that their model gives better fit than $G_1(1-p)$. They used Pearson χ^2 statistic and the discrepancy coefficient $C = \frac{\chi^2}{N}$ (N is the sample size) as the goodness of fit criterion (cf. Makčutek, (2008)) to support their claim. In the following, we shall use some more measures of goodness-of-fit to show that for this data set $GWG(\alpha,\beta,p)$ provides the best fit compared to $G_1(1-p)$, $WG(\alpha,p)$, $DGE(\gamma,p)$ and $DBE(\gamma,\lambda,p)$ models.

Table 4 shows the observed and expected values of the five models. The histogram of the data and fitted $GWG(\alpha,\beta,p)$, $WG(\alpha,p)$, $DGE(\gamma,p)$, $DBE(\gamma,\lambda,p)$ and $G_1(1-p)$ are plotted in Figure 8. Table 4 lists the MLEs of the parameters, Pearson χ^2 statistic, $C = \frac{\chi^2}{N}$, negative log-likelihood
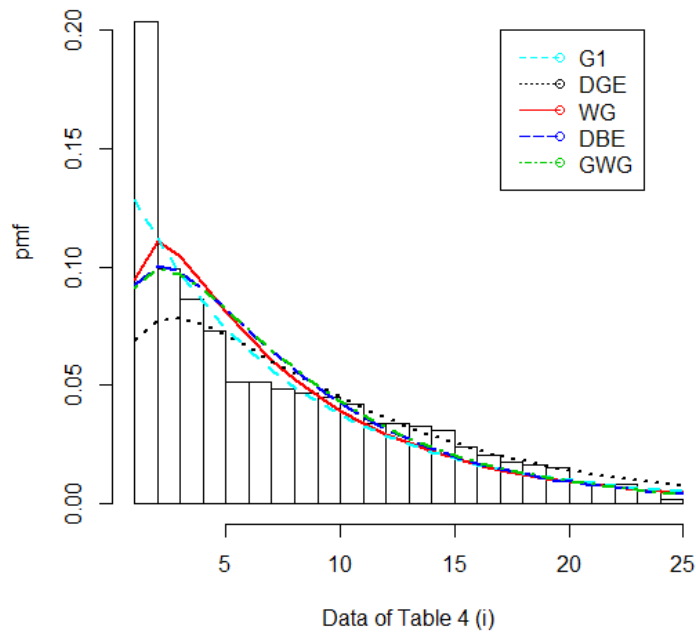
Fig. 8. The histogram of the data of Table 4 and the fitted pmfs of $G_1(1-p)$, $DGE(\gamma, p)$, $WG(\alpha, p)$, $DGE(\gamma, \lambda, p)$ and $GWG(\alpha, \beta, p)$.

function $(-\log L)$, Akaike Information Criterion, $AIC = -2\log L + 2k$, and Bayesian information criterion, $BIC = -2\log L + k\log(N)$, where k is the number of paramerets in the model and L is the maximized value of the likelihood function, for all five models.

Analysis of Table 5 indicates that by considering AIC and BIC indices, $DWE(\alpha, \beta, p)$ provides the best fit among the four models. Also, by comparing Pearson χ^2 statistic and the discrepancy coefficient $C = \frac{\chi^2}{N}$, we conclude that $GWG(\alpha, \beta, p)$ is the best fit among the five models. Since $DGE(\beta + 1, p)$ is a submodel of the class of $GWG(\alpha, \beta, p)$ distributions, to test $DGE(\beta + 1, p)$ versus $GWG(\alpha, \beta, p)$, we make the following test of hypotheses:

$$H_0 : \alpha = 1(DGE) \qquad v.s \qquad H_1 : \alpha \neq 1(GWG)$$

The value of the log likelihood ratio test (LLRT) statistic and its corresponding p-value are 23818.2 and 0, respectively. Therefore, the null hypothesis is rejected at any significance level.

Table 5. MLEs of the parameters, Pearson statistic, discrepancy coefficient, negative log-likelihood, AIC and BIC of the GWG(α, β, p), DBE(γ, λ, p), WG(α, p), DGE(γ, p) and $G_1(1-p)$ models for data of Table 4.

Distribution	$\hat{p}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\lambda}$	χ^2	C	$-\log L$	AIC	BIC
$GWG(\alpha, \beta, p)$	0.8325	0.0040	-	0.3908	-	12648.27	0.0403	932040.3	1864087	1864118
$DBE(\gamma, \lambda, p)$	0.8719	-	1.4004	-	1.2011	13008.59	0.0414	932580.0	1865166	1865236
$WG(\alpha, p)$	0.8637	7.0752	-	-	-	14398.35	0.0458	934695.4	1869395	1869416
$DGE(\gamma, p)$	0.8808	-	1.3950	-	-	14592.14	0.0465	943949.7	1887903	1887925
$G_1(1-p)$	0.8715	-	-	-	-	14846.44	0.0473	936366.9	1872738	1872759

7. Summary and Conclusion

In this article, a new generalization of weighted geometric (GWG) distribution is introduced whose special case can be viewed as a discrete analogue of generalized weighted exponential distribution. The proposed new distribution can also be constructed through a selection model. In addition, the class of GWG distribution enfolds several popular discrete distributions as special cases. Several important mathematical and structural properties of the GWG distribution are discussed. Our GWG distribution has provided the best fit among several rival models with one, two and three parameters for a real data set (cf. Makčutek (2008)).

References

[1] Y. Akdoğan, C. Kus, A. Asgharzadeh, I. Kinaci and F. Sharafi, Uniform-Geometric distribution, *Journal of Statistical Computation and Simulation,* **86**(9) (2016) 1754-1770 .

[2] A. Akinsete, F. Famoye and C. Lee, The Kumaraswamy-geometric distribution, *Journal of Statistical Distributions and Applications,* **1** (2014) 1-17.

[3] R. B. Arellano-Valle, M.D. Branco and M.G. Genton, A unified view on skewed distributions arising from selections, *The Canadian Journal of Statistics*, **34** (2006) 581-601.

[4] A. Azzalini, A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12** (1985) 171-178.

[5] M. Bebbington, C.D. Lai, M. Wellington and R. Zitikis, The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure data, *Reliability Engineering and System Safety*, **106** (2012) 37-44.

[6] H. Bidram, R. Roozegar and V. Nekoukhou, Exponentiated generalized geometric distribution: A new discrete distribution, *Hacettepe Journal of Mathematics and Statistics*, (2015) doi 10.

[7] M.G. Genton, *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*, Chapman and Hall/CRS, New York, (2004).

[8] E. Gómez-Déniz, Another generalization of the geometric distribution, *Test*, **19** (2010) 399-415.

[9] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*. seventh ed., Academic Press, San Diego.(2007).

[10] R.D. Gupta and D. Kundu, A new class of weighted exponential distributions, *Statistics*, **43** (2009) 621-634.

[11] S. S. Harandi and M. H. Alamatsaz, Discrete alpha-skew-Laplace distribution, *SORT: Statistics and Operations Research Transactions*, **39**(1) (2015) 71-84.

[12] S. Inusah and T. J. Kozubowski, A discrete analogue of the Laplace distribution, *Journal of Statistical Planning and Inference*, **136** (2006) 1090-1102.

[13] J. Keilson and H. Gerber, Some results for discrete unimodality, *Journal of the American Statistical Association*, **66** (1971) 386-389.

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

[15] T. J. Kozubowski and S. Inusah, A skew Laplace distribution on integers, *AISM*, **58** (2006) 555-571.

[16] J. H.C. Lisman and M. C. A. van Zuylen, Note on the generation of most probable frequency distributions, *Statistica Neerlandica*, **26** (1972) 19-23.

[17] J. Makčutek, A generalization of the geometric distribution and its application in quantitative linguistics, *Romanian Reports in Physics*, **60** (2008) 501-509.

[18] S. Nadarajah, The skew logistic distribution, *Advances in Statistical Analysis*, **93** (2009) 197-203 .

[19] V. Nekoukhou, M. H. Alamatsaz and H. Bidram, A discrete analogue of the generalized exponential distribution, *Communications in Statistics Theory Methods*, **41** (2011) 2000-2013.
    V. Nekoukhou, M. H. Alamatsaz, H. Bidram, and A. H. Aghajan, Discrete Beta-Exponential Distribution, *Communications in Statistics - Theory and Methods*, **44** (2015) 20792091.
[20] V. Nekoukhou and H. Bidram, Exponential-Discrete Generalized Exponential Distribution: A New Compound Model, *Journal of Statistical Theory and Applications*, **15.2** (2016) 169-180.
[21] V. Nekoukhou and H. Bidram, The exponentiated discrete Weibull distribution. *SORT* **39**(1) (2015) 127- 146.
[22] G.P. Patil and C.R. Rao, The weighted distributions: A survey and their applications, In: Krishnaiah, P.R. (Ed.), *Applications of statistics*, North-Holland Publ. Co., Amsterdam, (1977).
[23] G.P. Patil and C.R. Rao, Weighted distributions and size biased sampling with applications to wild-life populations and human families, *Biometrica*, **34** (1978) 179-189.
[24] A.N. Philippou, C. Georghiou and G.N. Philippou, A generalized geometric distribution and some of its properties, *Statistics Probability Letter*, **1** (1983) 171-175.
[25] M. Shaked and J.G. Shanthikumar, *Stochastic Orders*, Springer Verlag. New York, (2007).
[26] M, K. Shakhatreh, A two-parameter of weighted exponential distributions, *Statistics and Probability Letters*, **82** (2012) 252-261.
[27] F. W. Steutel and K. van Harn, *Infinite Divisibility of Probability Distributions on the Real Line*, New York: Marcel Dekker, (2004).
[28] R.C. Tripathi, R.C. Gupta and T.J. White, Some generalizations of the geometric distribution, *Sankhya Ser, B.*, **49** (1987) 218-223.