Unilag Journal of Mathematics and Applications,
Volume 4,Issue 1 (2024), Pages 51--69.
ISSN: 2805 3966. URL: http://lagjma.edu.ng

THE MODIFIED RANDOM WALK DISTRIBUTION

ADAMU LATEEF[1] AND FELIX FAMOYE

ABSTRACT: This study presents a modification to a two-parameter distribution, yielding a discrete distribution termed the modified random walk distribution. This modification is derived from the random walk distribution and falls within the family of Lagrangian probability distributions. Comprehensive investigations are conducted on both the random walk distribution and its modified counterpart to discern and analyze their respective properties. Furthermore, a recursive formula for generating probabilities associated with this distribution is proposed. This research contributes to the understanding of the random walk distribution and provides insights into the implications of its modification.

1. INTRODUCTION

A discrete probability distribution describes the probability of each value of a discrete random variable, which takes on distinct, countable values such as whole or natural numbers with gaps between them. Maity (2018) described discrete probability distribution as the probability of occurrence of each value of a discrete random variable. Classic distributions have historically played a pivotal role in modelling data across various domains, including engineering, actuarial science, environmental studies, medical sciences, biological research, demography, economics, finance, and insurance (Hossein, Pouya, & Ismail, 2016).

Given the fundamental role of discrete distributions in these diverse fields, researchers have contributed numerous methodologies to generate novel families of distributions. One noteworthy addition to this collection is the discrete Lagrangian probability distribution, named in honour of the eminent French mathematician Joseph-Louis Lagrange (1736-1813). Lagrange's profound contributions to mathematics encompassed analysis, number theory, mechanics, applied differential calculus, and probability theory. Notably, Lagrange co-created the calculus of variation, deriving the Euler-Lagrange equations, and introduced the method of solving differential equations known as the variation of parameters. His impact also extends to calculus, where he formulated the Lagrange remainder to the Taylor series and provided formulae for the expansion

2010 *Mathematics Subject Classification*. Primary: 22E30. Secondary: 58J05.
Keywords and phrases. Lagrange, Over-dispersion, Under-dispersion, Unimodal
©2024 Department of Mathematics, University of Lagos.
Submitted: August 9, 2024. Revised: September 23, 2024. Accepted: October 23, 2024.
[1]Correspondence

of a function f(z) in a power series of u when u = g(z). This paper delves into the discrete Lagrangian probability distribution and its implications for enhancing our understanding of various statistical phenomena.

Lagrange's expansion is the power series expansion of the inverse function of an analytic function, and it leads to general Lagrangian distributions (Imoto, 2016). The research work by Imoto (2016) presented some theorems in which different sets of two analytic functions provide a Lagrangian distribution.

Let $f(z)$ and $g(z)$ be two analytic functions of z, which are infinitely differentiable in $-1 \le z \le 1$ and such that $g(0) \ne 0$. Lagrange (1736-1813) considered the inversion of the Lagrange transformation $u = z/g(z)$, providing the value of z as a power series in u, and obtained

$$z = \sum_{k=1}^{\infty} \frac{u^k}{k!} \left[D^{k-1} \left(g(z) \right)^k \right]_{z=0,} \tag{1.1}$$

as power series expansion (Consul and Famoye 2006).

The Lagrange expansions can be used to obtain very useful numerous probability models. A number of discrete probability distributions are now available, of which several of these distributions and their properties have been discussed by Consul and Famoye (2006). Also, theorems and proofs which characterize the Lagrangian distributions are provided.

This paper presents an in-depth exploration of a discrete Lagrangian distribution, specifically identified as the random walk distribution. The analysis delves into the intricacies of this distribution, unravelling its unique characteristics and properties. Furthermore, the examination extends to encompass modifications of the random walk, shedding light on the implications and behaviours associated with such alterations. By systematically investigating these variants, valuable insights are provided into the broader understanding of discrete Lagrangian distributions and their nuanced applications. The comprehensive analysis presented enhances the knowledge of the random walk distribution and provides a foundation for exploring its modified forms and their implications in diverse fields.

2. MATERIALS AND METHODS

Lagrange Probability Distribution: The discrete Lagrangian probability distribution forms a very large and important class which contains numerous families of probability distributions (Consul & Famoye, 2006). The random walk distribution belongs to this class of distributions. The distribution is a product of the Lagrangian transformation of the function $z = ug(z)$. The random walk distribution belongs to the Delta Lagrangian distributions of the first kind, referred to as the Delta L_1 distribution.

According to Consul and Famoye (2006), the discrete Lagrangian probability distributions of the first kind (L_1) have been systematically studied in a number of papers by Consul and Shenton (1975). Several researchers have explored these areas of statistics and obtained various generalizations of some discrete distributions. This section aims to review several relevant probability distributions pertinent to this research.

2.1 Basic Lagrangian distribution

Consul and Famoye (2006) gave the definition: let the function $g(z)$ be a successively differentiable function such that $g(1)=1$ and $g(0)\neq 0$. The function $g(z)$ may or may not be a pgf. Then, the numerically smallest root $z=\ell(u)$ of the transformation $z=ug(z)$ defines a pgf $z=\psi(u)$ with the Lagrange expansion (1.1) in powers of u as

$$z=\psi(u)=\sum_{x=1}^{\infty}\frac{u^x}{x!}\left\{D^{x-1}\left(g(z)\right)^x\right\}_{z=0} \quad (2.1)$$

if $D^{x-1}\left(g(z)\right)^x\Big|_{z=0}\geq 0$ for all values of *x*.

Therefore, the class of basic Lagrangian distributions has probability mass function (pmf) as

$$P(X=x)=\frac{1}{x!}\left\{D^{x-1}\left(g(z)\right)^x\right\}_{z=0}, x\in\mathbb{N} \quad (2.2)$$

Some of the examples of the basic Lagrangian distributions are as follows:

Borel Distribution

The Borel distribution is a discrete probability distribution. It was named after the French mathematician Emile Borel. Tanner (1961) defined the distribution as a queuing process with random arrivals at rate q per unit time and constant service time β, the number of units served during a busy period.

The pmf is given by

$$P(x,\lambda)=\frac{e^{-x\lambda}\left(x\lambda\right)^{x-1}}{x!}, x=1,2,... \quad (2.3)$$

where $\lambda=\beta q$

The distribution in (2.3) was given by Borel (1942) and can be redefined by using the Lagrange's inversion formula. Selecting $g(z)=e^{\lambda(z-1)}, 0<\lambda<1$ in (2.2), the Borel distribution can be obtained. The mean and the variance are given by

$$\mu=\frac{\lambda}{(1-\lambda)} \text{ and } \sigma^2=\frac{\lambda}{(1-\lambda)^3} \text{ for } \lambda\in(0,1). \quad (2.4)$$

The Borel model satisfies the properties of under-dispersion and over-dispersion. The distribution is over-dispersed when λ satisfies the inequality $\frac{3}{2}-\sqrt{5}/2<\lambda<1$. It is under-dispersed when $\lambda<\frac{3}{2}-\sqrt{5}/2$ and is equal-dispersion when $\lambda=\frac{3}{2}-\sqrt{5}/2$.

Consul Distribution

The Consul distribution was introduced by Consul and Shenton (1975). A discrete random variable X is said to have Consul distribution if its pmf is given by

$$P\left(X=x\right)=\frac{1}{x}\binom{mx}{x-1}\theta^{x-1}\left(1-\theta\right)^{mx-x+1},\ x=1,2,... \tag{2.5}$$

By using $g\left(z\right)=\left(1-\theta+\theta z\right)^{m}$ in equation (2.2), the Consul distribution is obtained. The distribution in (2.5) reduces to the geometric distribution when m = 1.

The mean and variance of the distribution are

$$\mu=\frac{1}{\left(1-m\theta\right)} \text{ and } \sigma^{2}=\frac{m\theta\left(1-\theta\right)}{\left(1-m\theta\right)^{3}}. \tag{2.6}$$

The Consul distribution satisfies the dual properties of over- and under-dispersion. The model is under-dispersed for all values of $m\geq1$ when $\mu=\frac{\left(\sqrt{5}+1\right)}{2}$, and is over-dispersed for all values of $m\geq1$ when $\mu>2$ (Zahoor, Adil & Jan 2017). Islam and Consul (1990) modified the Consul distribution and derived it as a bunching model in traffic flow through branching process and also discussed its applications to automobile insurance claims.

Famoye (1997a) showed that the Consul distribution is the limit of zero-truncated generalized negative binomial distribution

$$P_{x}\left(\theta,\beta,m\right)=\frac{m}{m+\beta x}\binom{m+\beta x}{x}\theta^{x}\left(1-\theta\right)^{m+\beta x-x}/\left[1-\left(1-\theta\right)^{m}\right],\ x=1,2,3,... \tag{2.7}$$

as the parameter $\beta\rightarrow1$. The distribution is unimodal but not strongly unimodal for all values of $m\geq1$ and $0<\theta<1$ and the mode is at a point $x=1$.

Geeta Distribution

The Geeta distribution belongs to the basic Lagrangian distributions with pmf given by

$$P\left(X=x\right)=\frac{1}{\beta x+1}\binom{\beta x-1}{x}\theta^{x-1}\left(1-\theta^{\beta x-x}\right)\quad,x=1,2,... \tag{2.8}$$

The mean and variance of the Geeta distribution have been obtained as

$$\mu = \frac{(1-\theta)}{(1-\beta\theta)}, \quad \sigma^2 = \frac{(\beta-1)\theta(1-\theta)}{(1-\beta\theta)^3}. \tag{2.9}$$

The Geeta distribution satisfies the properties of over-dispersion and under-dispersion. Thus, for smaller values of θ, the Geeta distribution is under-dispersed and for larger values of $\theta < \beta^{-1}$ it is over-dispersed. The variance σ^2 equals the mean when $\theta = \frac{1}{2}\left[(3\beta-1)+\sqrt{(3\beta-1)^2+4\beta^2}\right]\beta^{-2}$ (Consul and Famoye, 2006). Other basic Lagrangian distributions include the Haight's distribution, Teja distribution, Felix distribution, and Sunil distribution.

2.2 Delta Lagrangian distribution

The delta Lagrangian distribution is another class of Lagrangian distributions and can be obtained from the basic Lagrangian distributions by taking the n-fold convolutions of the basic Lagrangian distributions. One of the methods of obtaining a delta Lagrangian distribution is to put $f(z) = z^n$ in the general Lagrangian expression

$f(z) = \sum_{k=1}^{\infty} a_k u^k$ with $a_k = \frac{1}{k!}\left[D^{k-1}(g(z))^k Df(z)\right]_{z=0}$ under the transformation $z = ug(z)$. The probability generating function (pgf) of the delta Lagrangian distribution becomes

$$z^n = (\psi(u))^n = \sum_{x=n}^{\infty} \frac{nu^x}{(x-n)!x}\left\{D^{x-n}(g(z))^x\right\}_{z=0}. \tag{2.10}$$

The probability mass function of the delta Lagrangian distribution is given as

$$P(X = x) = \frac{n}{(x-n)!x} D^{x-n}(g(z))^x |_{z=0}, \text{ for } x = n, n+1, n+2, ... \tag{2.11}$$

An example of the delta Lagrangian distributions is the Borel-Tanner distribution.

Borel-Tanner distribution

The Borel-Tanner distribution is an example of the delta Lagrangian distribution generalized from the Borel distribution. Tanner (1953) generalized the Borel distribution to show that the distribution of the number of units served in a busy period starting with an accumulation of *k* units. Tanner (1961) defined the generalized form as follows: Let *k* be a positive integer with parameter λ, then their sum $W = X_1 + X_2 + ... + X_k$ is said to have Borel-Tanner distribution.

The pgf is given by $e^{\lambda(z-1)}, 0 < \lambda < 1$ and the pmf is given as

$$P(W = n) = \frac{k}{n}\frac{e^{-\lambda n}(\lambda n)^{n-k}}{(n-k)!}, n = k, k+1, ... \tag{2.12}$$

The mean and variance of the Borel-Tanner distribution are

$$\mu = \frac{k}{1-\lambda} \text{ and } \sigma^2 = \frac{k\lambda}{(1-\lambda)^3} \quad (2.13)$$

The Borel-Tanner distribution has been used extensively in the analysis of data pertaining to biological word frequency, queuing and other branching areas (Gupta, & Jain 1977). In considering the applications, Gupta and Jain (1978) compare the bivariate distribution of Borel-Tanner with the bivariate negative binomial. Their work considers the probability distribution of the number of customers served in the busy period of a single server queue. Applications to traffic flow, semi-infinite discrete dam and branching processes are described. The result shows that in many situations the bivariate Borel-Tanner distribution provides a better fit than the bivariate negative binomial distribution.

Other examples of the delta Lagrangian distributions are delta binomial distribution, delta Geeta distribution, random walk distribution, delta Katz distribution, delta-Felix, and delta-Teja distributions.

2.3 General Lagrangian distributions

Consul and Famoye (2006) gave the definitions of the general Lagrangian distributions as follows: Let $g(z)$ and $f(z)$ be two analytic functions of *z* differentiable with respect to *z*, such that $g(0) \neq 0$, $g(1)=1$, $f(1)=1$ and

$$D^{x-1}\left\{\left(g(z)\right)^x f'(z)\right\}|_{z=0} \geq 0 \text{ for} x \in \mathbb{N}. \quad (2.14)$$

The pgf of the discrete general Lagrangian probability distribution, under the Lagrangian transformation $z = ug(z)$, is given by $f(z) = \sum_{k=o}^{\infty} a_k u^k$ in the form

$$f(z) = f(\psi(u)) = \sum_{x=0}^{\infty}\left(\frac{u^x}{x!}\right)D^{x-1}\left\{\left(g(z)\right)^x f'(z)\right\}|_{z=0}, \quad (2.15)$$

The pmf of the general Lagrangian distributions becomes

$$P(X = x) = \left(\frac{1}{x!}\right) D^{x-1}\left\{\left(g(z)\right)^x f'(z)\right\}|_{z=0}, x \in \mathbb{N}. \quad (2.16)$$

Some examples of the general Lagrangian distributions are Binomial Poisson, Poisson-binomial, generalized binomial, generalized Katz, modified Felix, and Shenton distributions.

3. RESULT

Result Random Walk Distribution: A random walk is a mathematical object, known as a stochastic or random process, which describes a path consisting of a succession of random steps on a mathematical space such as an integer. Lawler and Limic (2010) defined random walk as a stochastic process formed by successive summation of independent, identically distributed random variables. The random walk is one of probability theory's most basic and well-studied topics. Consul and Famoye (2006) defined the random walk as a particle which is said to perform a simple random walk on a line when starting from an initial position '*n'* (where *n* is an integer) on the line; it moves each time from its position either a unit step (+1) in a positive direction with some probability '*p*' or a unit step (-1) in the negative direction with probability $q = 1 - p$. The walk is a discrete-time, which is a sequence of random variables. Various probabilities are assigned for the various possible mutually exclusive moves or jumps of the particle such that the sum of the probabilities is one. The average total displacement is zero. Random walks explain the observed behaviours of many processes and can serve as a model for stochastic activities. According to Xian et al. (2019), Random walks can be utilized to analyze and simulate the randomness of objects and to calculate the correlation among them, making this method useful for solving practical problems.

Examples of random walks or random processes are the movement of particles, the motion of molecules in a solution, stock market fluctuations, electron diffusion in metals, fluctuations in crude oil prices, etc.

Researchers who work with particle systems and other models that use random walks as a basic ingredient often need more precise information on random walk behaviour than that provided by the central limit theorems (Lawler & Limic, 2010). Random walks have many applications in fields such as ecology, computer science, physics, chemistry, economics, etc. Xia et al. (2019) provided reviews of classical random walks and quantum walks and introduced their applications in the field of computer science. Bhattacharyya et al. (2023) introduced the fundamental features of the simple random walk problem in one dimension, providing both analytical and numerical analyses of symmetric and asymmetric random walks. They also explored the probability of a walker returning to the starting point for the first time. There are various types of random walks and have been extensively studied in the literature, with significant contributions made to understanding their properties and applications. Notable works include those by Jain and Orey (1973), Weiss and Havlin (1986), Hills (2020), and Ahsanullah et al. (2023). Hills (2020) provide a comprehensive analysis of the statistical properties and theorems of various classes of random walks, using both analytical and numerical approaches. It also explores the applications of random walks and discusses different methods of analysis, beginning with an introduction to the types of random walks and their properties. However, the term random walk often refers to a special category of Markov chains or Markov processes, but many time-dependent processes are referred to as random walks. The number of different walks of n-steps where each step is $+1$ or -1 is 2^n. For the simple random walk, each of these walks are equally likely.

3.1 Definitions and Properties of the Random Walk Distribution

Among the Delta Lagrangian (L_1) distributions is the random walk distribution. The random walk distribution is obtained by taking the n-fold convolution of the probability generating function (pgf) $g(z)= p+qz^2$. The probability mass function (pmf) of the delta Lagrangian distribution is written as $P(X=x)=\frac{m}{(x-m)!x}D^{x-m}\left[g(z)\right]^x|_{z=0}$, (3.1)

where $g(z)=p+qz^2$, and substituting this into (3.1), we have

$$P(X=x)=\frac{m}{(x-m)!x}D^{x-m}\left[p+qz^2\right]^x|_{z=0}.$$

By simplifying the above, we have

$$P(X=x)=\frac{m}{(x-m)!x}D^{x-m}\sum_{k=0}^{x}\binom{x}{k}p^{x-k}\left(qz^2\right)^k|_{z=0}$$

$$=\frac{m}{(x-m)!x}D^{x-m}\sum_{k=0}^{x}\binom{x}{k}p^{x-k}q^k z^{2k}|_{z=0}$$

$$=\frac{m}{(x-m)!x}\sum_{k=0}^{x}\binom{x}{k}p^{x-k}q^k D^{x-m}z^{2k}|_{z=0}$$

$$=\frac{m}{(x-m)!x}\sum_{k=0}^{x}\binom{x}{k}p^{x-k}q^k\,2k(2k-1)(2k-2)\ldots(2k-(x-m)+1). \quad (3.2)$$

But $2k=x-m$, so that the power of z is zero in the expansion.

So $k=(x-m)/2$. (3.3)

Substituting (3.3) into (3.2), we have

$$P(X=x)=\frac{m}{(x-m)!x}\binom{x}{(x-m)/2}p^{x-(x-m)/2}q^{(x-m)/2}(x-m)!$$

$$=\frac{m}{(x-m)!x}\frac{x!}{(x-(x-m)/2)!((x-m)/2)!}p^{x-(x-m)/2}q^{(x-m)/2}(x-m)!$$

$$=\frac{m}{x}\binom{x}{(x-m)/2}p^{(x+m)/2}q^{(x-m)/2}$$

Since $(m+x)/2=(2m-m+x)/2=m+(x-m)/2$

$$P(X=x)=\frac{m}{x}\binom{x}{(x-m)/2}p^m p^{(x-m)/2}q^{(x-m)/2}$$

$$= \frac{m}{x}\binom{x}{(x-m)/2} p^m (pq)^{(x-m)/2}, \qquad x = m, m+2, m+4, \ldots \tag{3.4}$$

Equation (3.4) is the random walk distribution.

3.2 Factorial Moments

To derive the factorial moment, we introduce the Lagrangian transformation $z = ug(z)$ and find the derivative with respect to *u*. In this case, the pgf of the distribution is a power series in *u* given by

$$f(z) = f(\psi(u)), \tag{3.5}$$

where $z = \psi(u)$. To obtain the factorial moment, we differentiate (3.5) with respect to u, then set $u = z = 1$.

$$f(z) = f(ug(z))$$

$$\frac{\partial f(z)}{\partial u} = \frac{\partial f(ug(z))}{\partial u}$$

$$= f'(z)g(z) + ug'(z)\frac{\partial f(z)}{\partial u}.$$

$$\frac{\partial f(z)}{\partial u} - ug'(z)\frac{\partial f(z)}{\partial u} = f'(z)g(z)$$

$$\frac{\partial f(z)}{\partial u}(1 - ug'(z)) = f'(z)g(z)$$

$$\frac{\partial f(z)}{\partial u} = \frac{f'(z)g(z)}{1 - ug'(z)} \tag{3.6}$$

$$\mu'_{(1)} = \left.\frac{f'(z)g(z)}{[1 - ug'(z)]}\right|_{u=z=1} = \frac{f'(1)g(1)}{1 - g'(1)} = \frac{f'}{1 - g'}. \tag{3.7}$$

Equation (3.7) is the first factorial moment $\mu'_{(1)}$. For the second factorial moment, we have to differentiate (3.6) with respect to u,

$$\frac{\partial^2 f(z)}{\partial u^2} = \frac{\partial}{\partial u}\left[f'(z)g(z) + ug'(z)\frac{\partial f(z)}{\partial u}\right]$$

$$= f''(z)\frac{dz}{du}g(z) + f'(z)g'(z)\frac{dz}{du} + \left[g'(z) + ug''(z)\frac{dz}{du}\right]\frac{\partial f(z)}{\partial u} + ug'(z)\frac{\partial^2 f(z)}{\partial u^2} \tag{3.8}$$

$$\frac{\partial^2 f(z)}{\partial u^2} - ug'(z)\frac{\partial^2 f(z)}{\partial u^2} = f''(z)\frac{dz}{du}g(z) + f'(z)g'(z)\frac{dz}{du} + \left[g'(z) + ug''(z)\frac{dz}{du}\right]\frac{\partial f(z)}{\partial u} \quad (3.9)$$

But, $z = ug(z)$

$$\frac{dz}{du} = \frac{g(z)}{1-ug'(z)} \quad (3.10)$$

Substituting (3.6) and (3.10) into (3.9), we have,

$$\frac{\partial^2 f(z)}{\partial u^2}\left(1-ug'(z)\right) = \frac{f''(z)g(z)g(z)}{1-ug'(z)} + \frac{f'(z)g'(z)g(z)}{1-ug'(z)} + \left[g'(z) + ug''(z)\frac{dz}{du}\right]\frac{f'(z)g(z)}{1-ug'(z)}$$

$$= \frac{\left[f''(z)g(z) + f'(z)g'(z)\right]g(z)}{1-ug'(z)} + \frac{f'(z)g(z)}{1-ug'(z)}\left[g'(z) + ug''(z)\frac{dz}{du}\right]$$

$$\frac{\partial^2 f(z)}{\partial u^2} = \frac{\left[f''(z)g(z) + f'(z)g'(z)\right]g(z)}{\left[1-ug'(z)\right]^2} + \frac{f'(z)g(z)}{\left[1-ug'(z)\right]^2}\left[g'(z) + ug''(z)\frac{dz}{du}\right]$$

$$\mu'_{(2)} = \left.\frac{\left[f''(z)g(z) + f'(z)g'(z)\right]g(z)}{\left[1-ug'(z)\right]^2} + \frac{f'(z)g(z)}{\left[1-ug'(z)\right]^2}\left[g'(z) + ug''(z)\frac{dz}{du}\right]\right|_{u=z=1}$$

$$\mu'_{(2)} = \frac{f''+2f'g'}{(1-g')^2} + \frac{f'g''}{(1-g')^3} \quad (3.11)$$

Recall that,

$f(z) = z^m$, $f'(z) = mz^{m-1}$, $f''(z) = m(m-1)z^{m-2}$. Hence,

$$f = 1,\ f'(1) = f' = m, \text{ and } f'' = m(m\text{-}1). \quad (3.12)$$

$g(z) = p + qz^2$, $g'(z) = 2qz$, $g''(z) = 2q$. Hence,

$$g = 1,\ g' = 2q \text{ and } g'' = 2q. \quad (3.13)$$

The above functions will be substituted where necessary to obtain the mean and variance.

3.3 Mean and Variance of Random Walk

The mean and variance of the random walk distribution in equation (3.4), can be obtained from the first and second factorial moments derived in equations (3.7) and (3.11).

Mean

From the first moment, $\mu'_{(1)} = \left.\frac{f'(z)g(z)}{\left[1-ug'(z)\right]}\right|_{u=z=1}$

$$\mu = \mu'_{(1)} = \frac{f'}{1-g'} = \frac{m}{1-2q}, \tag{3.14}$$

by using (3.12) and (3.13).

The mean of the random walk distribution is represented by equation (3.14).

Variance

We know that $\mu'_{(2)} = E[X(X-1)] = E(X^2) - E(X)$ and $\sigma^2 = E(X^2) - \mu^2$. Hence,

$$\sigma^2 = \mu'_{(2)} + \mu - \mu^2 \tag{3.15}$$

By using equations (3.11), (3.12) and (3.13), we have

$$\sigma^2 = \frac{f''+2f'g'}{(1-g')^2} + \frac{f'g''}{(1-g')^3} + \frac{f'}{1-g'} - \frac{(f')^2}{(1-g')^2}$$

$$= \frac{f''+2f'g'+(f'-f'g')-(f')^2}{(1-g')^2} + \frac{f'g''}{(1-g')^3}$$

$$= \frac{f''+f'g'+f'-(f')^2}{(1-g')^2} + \frac{f'g''}{(1-g')^3}$$

By using (3.12) and (3.13), we have

$$\sigma^2 = \frac{m(m-1)+2mq+m-m^2}{(1-2q)^2} + \frac{2mq}{(1-2q)^3}$$

$$= \frac{m^2-m+2mq+m-m^2}{(1-2q)^2} + \frac{2mq}{(1-2q)^3}$$

$$= \frac{2mq}{(1-2q)^2} + \frac{2mq}{(1-2q)^3} = \frac{2mq(1-2q)+2mq}{(1-2q)^3} = \frac{4mq-4mq^2}{(1-2q)^3}$$

$= \frac{4mq(1-q)}{(1-2q)^3}$, $q < 1/2$. Hence,

$$\sigma^2 = 4mq(1-q)(1-2q)^{-3} \tag{3.16}$$

3.4 Dispersion Property for Random Walk Distribution

In explaining the dispersion properties inherent to the random walk distribution, a comprehensive analysis requires the investigation of the variance-to-mean ratio. This critical examination not only quantifies the variability within the distribution but also contributes valuable insights into the stochastic dynamics governing the random walk process. By delving into the interplay between variance and mean, the aim is to unveil nuanced aspects of dispersion that hold significance in various applications, enhancing the understanding of the underlying statistical behavior and implications of the random walk distribution. That is, $\sigma^2/\mu = \left(\frac{4mq(1-q)}{(1-2q)^3}\right)\Big/\left(\frac{m}{1-2q}\right)$. If $\sigma^2/\mu > 1$, the distribution is said to be over-dispersed, if $\sigma^2/\mu < 1$

It is said to be under-dispersed and it is equi-dispersed when $\sigma^2/\mu = 1$. However, it is necessary to solve for the value of the parameter *q*. To achieve this, $\sigma^2 > \mu$, is considered, based on the values of the mean and variance in equations (3.14) and (3.16), respectively.

$$\sigma^2 > \mu \text{ when } \frac{4mq(1-q)}{(1-2q)^3} > \frac{m}{1-2q}$$

$$\therefore \quad \frac{4q(1-q)}{(1-2q)^2} > 1$$

$$4q - 4q^2 > 1 - 4q + 4q^2$$

$$8q^2 - 8q + 1 < 0$$

Solving the quadratic function, the value of *q* will be,

$$q = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{8 \pm \sqrt{64 - 32}}{16}$$

$$q = 1/2 \pm \sqrt{2}/4$$

Since $q < 1/2$, then the required value of *q* will be $1/2 - \sqrt{2}/4$. It is also necessary to determine the region of parameter *q* that makes the distribution either over-, under-, or equi-dispersed. The following regions for the parameter *q*: $0 < q < 1/2 - \sqrt{2}/4$ and $1/2 - \sqrt{2}/4 < q < 1/2$, will be considered, and the region where the quantity $8q^2 - 8q + 1$ is negative will be identified. Since $q < 1/2$, the following conditions apply:

The distribution is over-dispersed when $1/2 - \sqrt{2}/4 < q < 1/2$.

The distribution is under-dispersed when $0 < q < 1/2 - \sqrt{2}/4$.

The distribution is equi-dispersed when $q = 1/2 - \sqrt{2}/4$.

4. DISCUSSION

4.0 Modified Random Walk Distribution: The Modified Random Walk (MRW) distribution is an extension derived from the foundational Random Walk distribution, as outlined in (3.4). The latter is expressed as:

$$P(X = x) = \frac{m}{x}\binom{x}{(x-m)/2} p^m (pq)^{(x-m)/2}, \quad x = m, m+2, m+4, \dots$$

We modify (3.4) by setting

$$y = (x-m)/2, \tag{4.0}$$

such that $x = 2y + m$. Then, the MRW distribution becomes

$$P(Y = y) = \frac{m}{m+2y}\binom{m+2y}{y} p^m (pq)^y$$

$$= \frac{m}{m+2y}\binom{m+2y}{y} p^{m+y} q^y, \quad y = 0,1,2,\dots \tag{4.1}$$

Therefore, equation (4.1) is referred to as the modified random walk distribution.

4.1 **Recursive formula for Computing the Probability of MRW Distribution**

A recursive formula is employed to compute the probabilities associated with the Modified Random Walk (MRW) distribution. Initially, the probability at $y = 0$ is computed, serving as the initial term to define the subsequent terms in the sequence. Utilizing the probability mass function (PMF) of the modified random walk, as presented in (3.21), and setting $y = 0$, the following is obtain:

$$P(Y = y) = \frac{m}{m+2y}\binom{m+2y}{y}(1-q)^{m+y} q^y \Big|_{y=0}.$$

The probability is given by

$$P(Y = 0) = P_0 = \frac{m}{m+0}\binom{m+0}{0}(1-q)^{m+0} q^0 = (1-q)^m$$

The recursive relationship is defined by the following:

$$P(Y=y)=P_y=\frac{m}{m+2y}\binom{m+2y}{y}(1-q)^{m+y}\,q^y$$ and

$$P_{y-1}=\frac{m}{m+2(y-1)}\binom{m+2(y-1)}{y-1}(1-q)^{m+(y-1)}\,q^{y-1}$$

$$\frac{P_y}{P_{y-1}}=\frac{\frac{m}{m+2y}\binom{m+2y}{y}(1-q)^{m+y}\,q^y}{\frac{m}{m+2(y-1)}\binom{m+2(y-1)}{y-1}(1-q)^{m+(y-1)}\,q^{y-1}}$$

$$=\frac{\frac{1}{m+2y}\binom{m+2y}{y}(1-q)^{m+y}\,q^y}{\frac{1}{m+2y-2}\binom{m+2y-2}{y-1}(1-q)^{m+y-1}\,q^{y-1}}$$

$$=\frac{\frac{1}{m+2y}\binom{m+2y}{y}(1-q)^{m+y-(m+y-1)}\,q^{y-y+1}}{\frac{1}{m+2y-2}\binom{m+2y-2}{y-1}}$$

$$=\frac{\frac{1}{m+2y}\frac{(m+2y)!}{(m+y)!\,y!}}{\frac{1}{m+2y-2}\frac{(m+2y-2)!}{(m+y-1)!(y-1)!}}(1-q)q$$

$$=\frac{(m+2y)!}{(m+2y)(m+y)!\,y!}\frac{(m+2y-2)(m+y-1)!(y-1)!}{(m+2y-2)!}(1-q)q$$

$$=\frac{(m+2y)(m+2y-1)(m+2y-2)!}{(m+2y)(m+y)(m+y-1)!\,y(y-1)!}\frac{(m+2y-2)(m+y-1)!(y-1)!}{(m+2y-2)!}(1-q)q$$

$$=\frac{(m+2y-1)(m+2y-2)}{(m+y)y}(1-q)q$$

$$P_y=\frac{(1-q)q(m+2y-1)(m+2y-2)P_{y-1}}{(m+y)y},\ y=1,2,3,\ldots$$

For $y=1$,

$$P_1 = \frac{(1-q)q(m+2-1)(m+2-2)P_0}{(m+1)} = mP_0(1-q)q$$

Recall that $P_0 = (1-q)^m$, then $P_1 = m(1-q)^m (1-q)q$.

To extend this computation to other values of y, a consistent methodology will be applied. In the course of this research work, the probabilities will be iteratively computed, leveraging the capabilities of the R statistical package for efficient and rigorous analysis.

Graphical Representation of the Modified Random Walk (MRW) Distribution

The MRW distribution is visually explored through Figures 1 to 4, where various aspects of its behaviour are elucidated. In order to investigate the impact of the dispersion parameter on probabilities, the parameter is systematically varied, and its influence on the distribution is observed.

A noteworthy characteristic observed in these graphs is the directional skewness, notably skewed to the right. Furthermore, the distribution is unimodal, exhibiting a single peak. In Figure 1, the peak is situated around zero and appears to shift to the right as the parameter m increases. Concurrently, the spread around the peak expands with an increase in m. Figure 4, characterized by the highest m value among the figures, and depicts a noteworthy phenomenon— the peak gravitates toward the center (mean) as the parameter escalates.

These visualizations not only provide a comprehensive understanding of the modified random walk distribution but also offer valuable insights into the distribution's behavior under varying dispersion parameters.

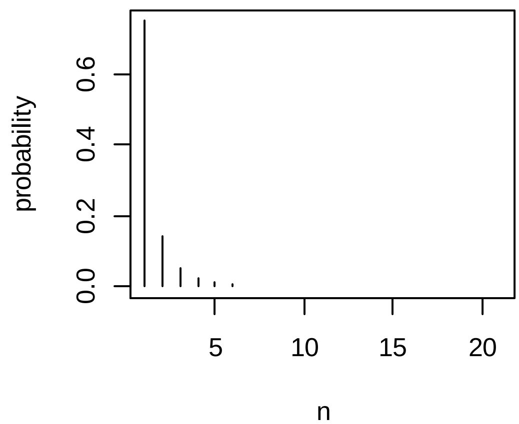

Figure 1. m=1

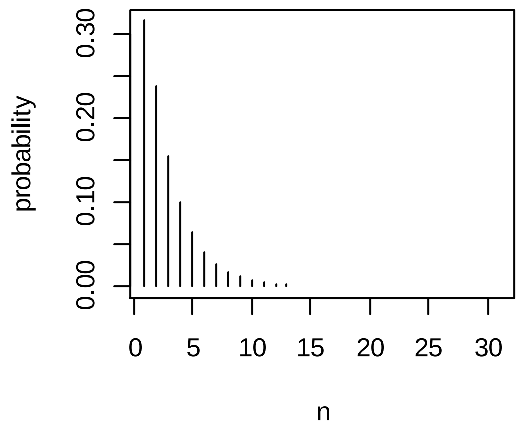

Figure 2. m=4

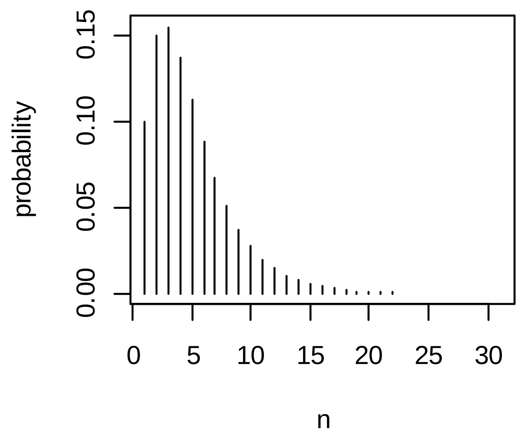

Figure 3. m=8

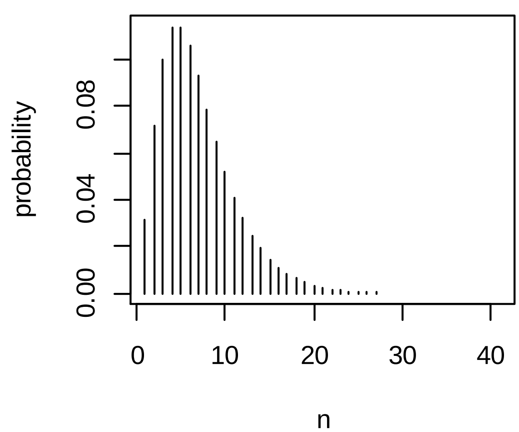

Figure 4. m=12

4.2 Mean and Variance of Modified Random Walk Distribution

The mean and variance of the MRW distribution will be obtained from that of the random walk distribution in equations (3.14) and (3.16). This is done by taking the expectation of (4.0). In doing so, the following is obtained:

$$E(Y) = \mu = [E(X) - m]/2 = mq/(1-2q), \; q < 1/2, \tag{4.2}$$

by using equation (4.0). Therefore the mean for the MRW distribution is given by (4.2).

Variance

From Equation (4.0), we have

$$V(Y) = \sigma^2 = V(X)/4 = mq(1-q)/(1-2q)^3, \; q < 1/2. \tag{4.3}$$

Also, the variance for the MRW distribution is given by (4.3).

4.3 Dispersion Property for MRW Distribution

To study the dispersion property of the MRW distribution, it is necessary to examine the relationship between the mean and the variance—whether they are equal or one is greater than the other. This is done by considering:

$$\sigma^2 > \mu \text{ or } \frac{mq(1-q)}{(1-2q)^3} > \frac{mq}{1-2q} \tag{4.4}$$

On simplification, (4.4) gives

$$1-q > 1-4q+4q^2 \Rightarrow -4q^2+3q > 0 \Rightarrow q(-4q+3) > 0 \text{ or } q(q-3/4) < 0.$$

Since $0 < q < 0.5$, the inequality $q(q-3/4) < 0$ is always satisfied. Hence, the MRW distribution is always over-dispersed.

5. CONCLUSION

This study focused on the Modified Random Walk (MRW) distribution, a discrete probability distribution derived from the random walk distribution. The primary objective was to enhance the understanding of the random walk distribution through modification and an in-depth exploration of its properties. The MRW distribution, characterized by the parameters m and q, was subject to comprehensive analysis, including the derivation of key properties such as the mean and variance, with particular emphasis on studying its dispersion characteristics.

A recursive formula for generating distribution probabilities was developed, and iterative computations of probabilities were performed using the R statistical package. Additionally, graphical representations for various dispersion parameters were presented. The findings indicate that the dispersion property of the MRW distribution consistently demonstrates over-dispersion, where the variance consistently exceeds the mean.

The computed probabilities, as illustrated in the accompanying graphs, lead to the conclusion that the MRW distribution is unimodal and consistently exhibits right skewness. These findings provide valuable insights into the behaviour and characteristics of the distribution, highlighting its statistical properties and potential applications.

Acknowledgment. The authors express their gratitude to the anonymous reviewers for their useful comments and suggestions.

Authors Contributions. Each author contributed equally to the work.

Authors' Conflicts of Interest. None

Funding Statement. There is no funding statement.

REFERENCES:

[1] M. Ahsanullah, V.B. Nevzorov, & A.V. Stepanov. Distributions and Characterizations Associated with a Random Walk. *J Math Sci* 273, (2023). 667–675 https://doi.org/10.1007/s10958-023-06531-w

[2] A. Bhattacharyya, J. Mitra, S. Bhattacharyya. A Brief Study on Simple Random Walk in 1D. *Reson* 28, (2023) 945–958 https://doi.org/10.1007/s12045-023-1624-2

[3] E. Borel. Surl' employ du theorem de Bernoulli pour faci liter, le calcul d' un infinite de coefficients. Application un problem de l' attente a un guichet. *Comptes Rendus Academie des Sciences, Paris Series A*, (214), (1942) 452-456.

[4] P. C. Consul, F. Famoye. Lagrangian Probability Distributions. Birkhauser, Boston, USA, (2006).

[5] P.C. Consul, L.R. Shenton. On the probabilistic structure and properties of discrete Lagrangian distributions. A modern course on Statistical Distributions in Scientific work, D. Riedel, Boston, M.A1, (1975) 41-57.

[6] F. Famoye. Generalized geometric distribution and some applications. Journal of Mathematical Sciences (8), (1997a) 1-13.

[7] R. P. Gupta, G. C. Jain. A bivariate Borel-Tanner distribution and its approximations. Journal of Statistical Computation and Simulation. 6(2) (1977) 129-135.

[8] R.P. Gupta, G.C. Jain. Some applications of a bivariate Borel-Tanner distribution. Biometrical Journal. 20(2) (1978) 99-105.

[9] A. Hills. Random walks: The Properties, Applications and Methods of Analysis. Bachelor's Thesis, University of Groningen, Groningen Netherlands (2020).

[10] Z. Hossein, F. Pouya, N. Ismail. Bivariate generalized Poisson regression model: Applications on health care data. Empirical Economics, Journal of the Institute for Advanced Studies. 51(4), (2016) 1607-1621.

[11] T. Imoto. Properties of Lagrangian distribution. Communications in Statistics-Theory and Methods, 45(3) (2016) 712-721.

[12] M. N. Islam, P. C Consul. The Consul distribution as a bunching model in traffic flow. Transportation Research Part B: Methodological. (25), (1990) 365-372.

[13] N. C. Jain, & S. Orey, Some properties of random walk paths. Journal of Mathematical Analysis and Applications, 43(3), (1973) 795– 815. doi:10.1016/0022- 247x(73)90293-x

[14] G. F. Lawler, V. Limic. Random Walk: A Modern Introduction (Cambridge studies in Advance Mathematics) Cambridge University Press, Cambridge, England (2010).

[15] R. Maity. Probability Distributions and Their Applications. Springer Transactions in Civil and Environmental Engineering, (2018). *93–143.* doi:10.1007/978-981-10-8779-0_4

[16] J. C. Tanner. A derivation of the Borel distribution. Biometrika, (48), (1961) 222-224.

[17] J. C. Tanner. A problem of interference between two queues. *Biometrika*, (40), (1953) 58-69.

[18] G. H. Weiss, S. Havlin. Some properties of a random walk on a comb structure, Physica A: Statistical Mechanics and its Applications, Volume 134, Issue 2, (1986) Pages 474-482

[19] F. Xia, J. Liu, H. Nie, Y. Fu, L. Wan, X. Kong. Random Walks: A Review of Algorithms and Applications. IEEE Transactions on Emerging Topics in Computational Intelligence, (2019). 1–13. doi:10.1109/tetci.2019.2952908

[20] A. Zahoor, R. Adil, T. R. Jan. A new discrete compound distribution with application. Journal of Statistics Applications and Probability 6(1) (2017) 233-241.

ADAMU LATEEF*
DEPARTMENT OF STATISTICS, FACULTY OF SCIENCE, UNIVERSITY OF LAGOS, AKOKA, NIGERIA.
Email Address: adamu.lateef1@gmial.com

FELIX FAMOYE
DEPARTMENT OF STATISTICS, CENTRAL MICHIGAN UNIVERSITY MT. PLEASANT, MICHIGAN, USA.
Email Address: felix.famoye@cmich.edu