Modeling Discrete Life Data in Reliability and its Applications

By

Najla Saad Rashid Alhazzani

Submitted in Partial Fulfillment of the Requirements
For
Master Degree of Science in Statistics

Department of Statistics and Operations Research
College of Science
King Saud University
March 2012

Kingdom of Saudi Arabia
King Saud University
College of Science
Statistics & Operations Research Department

The undersigned hereby certify that, they have read and recommend to the College of Science for acceptance a thesis entitled **"Modeling Discrete Life Data in Reliability and its Applications"**, by **Najla Saad Rashid Alhazzani** in partial fulfillment of the requirements for the degree of Master of Science.

Dated: March, 2012

Name **signature**

Principal Supervisor: Prof. Abdulrahman M. Abouammoh

Co- Supervisor: Dr. Arwa M. A. Alshingiti

Committee Members:

Dr. Amna Abdulatif Jamjoom

Prof. Hicham Mustafa Al-Nachawati

Prof. Mohamed Ataharul Islam

Abstract

Almost all reliability studies assume models with underlying continuous variables. However, discrete data sets are shown in many practical situations. For example, systems functioning might be measured in cycles, runs or sometimes monitored once per period. In such cases, the lifetimes are the number of functioning cycles or the number of time periods successfully completed before the failure. In addition, reliability data are often grouped or rounded that are treated as a discrete data. For these reasons discrete statistical modelling of reliability data and variables is indispensible area of research in theory of reliability and life testing.

In this thesis, discrete counter-part of two important continuous distributions, namely, gamma and Lindley are introduced. In fact, two types of the expected discrete Lindley that are derived by mixing gamma distributions with different shape parameters are established. The statistical and reliability properties of each distribution are derived. Some interesting interrelationships between various discrete distributions are also explored. Furthermore, the size-biased versions of some of these distributions are derived and discussed. Different estimation methods of the underlying parameters for these distributions are utilized and comparisons of their performance have been made.

Table of Contents

List of Figures

List of Tables

Notations

pmf	probability mass function
pdf	probability density function
cdf	cumulative distribution function
ecdf	empirical cumulative distribution function
sf	survival function
FR	failure rate
MRL	mean remaining life
DGD(α,q)	discrete gama distribution
CGD(α,β)	continuous gamma distribution
GDLD(α,q)	generalized discrete Lindley distribution
$GDLD_2(2,q)$	generalized discrete Lindley distribution II
G(q)	geometric distribution
NB(r,q)	negative binomial distribution
WI(α,q)	type I discrete Weibull distribution
WII(α,m)	type II discrete Weibull distribution
WIII(α,c)	type III discrete Weibull distribution
PL(θ)	Poisson-Lindley distribution
LBPL(θ)	length-biased Poisson-Lindley distribution

NBL(r,θ)	negative binomial-Lindley distribution
mgf	moment generating function
IFR	increasing failure rate
CFR	constant failure rate
DFR	decreasing failure rate
IFRA (DFRA)	increasing (decreasing) failure rate in average
NBU (NWU)	new better (worse) than used
NBUE (NWUE)	new better (worse) than used in expectation
mle	maximum likelihood estimator
mse	mean square error
MVUE	minimum variance unbiased estimator
CR	Cramer-Rao

Chapter 1

Introduction

This chapter presents an overview of historical background and main statistical and reliability concepts that are used in the rest of this thesis. It also discusses briefly properties of estimators along with some estimation methods. Moreover, some goodness-of-fit techniques are also indicated. Finally, the scope of the thesis is fully described.

1.1 Historical background

In the past 50 years, remarkable works have been carried out in the field of reliability. Almost all published studies in reliability usually consider continuous reliability data and few works have been done for discrete reliability models. Further, limited number of discrete lifetime distributions have been proposed. An account of the most commonly used distributions is given in Johnson et al. (1992). They have also explored interrelationships between different distributions. In this thesis we consider mainly discrete reliability models. This is supported by the fact that discrete data sets are the most commonly observable in real situations. One also notes that components and systems have natural discrete measures for life times such as functioning in cycles, time units, or those monitored once per period. In such situations, the lifetimes are expressed in terms of the number of functioning cycles or the number of time periods successfully completed before failure. Moreover, the reliability data might be available to the investigator in discrete measures

because they are grouped or they are rounded, meaning the exact measurement is within an interval.

Several discrete lifetime distributions have been proposed in the literature. The well-known group of the discrete distributions includes geometric, binomial, negative binomial, Poisson … etc. There are also discrete distributions defined based on their continuous counter-parts. For instance, Nakagawa and Osaki (1975) have proposed the first discrete counter-part of the usual continuous Weibull distribution. Their proposed form is based on the similarity of survival functions for discrete and continuous Weibull. Basic results about discrete reliability have been presented by Salvia and Bollinger (1982) where they illustrated them by a simple discrete distribution with only one parameter. The second discrete Weibull distribution was proposed by Stein and Dattero (1984) and it is derived from the similarity of the failure rates for discrete and continuous Weibull. The third one was proposed by Padgett and Spurrier (1985). These distributions have not received much attention by researchers until estimation of parameters by the method of proportion for these distributions are discussed in Ali Khan et al. (1989).

Nowadays, with the huge growth in the collection and storage of data due to technological advances, count data have become immensely available in many fields such as transportation safety and clinical research. Examples include the number of crashes in highways, the number of visits to a website and the number of calls to a call center. The most popular distribution for modeling count data has been the Poisson distribution. Although Poisson models are very popular for modeling count data, many modelling cases in financial and actuarial science do not adhere to the assumption of equi-dispersion where the mean and the variance are equal. One way of dealing with such a case is by using a

distribution with more than one parameter. A common used method of obtaining an additional parameter in the distribution is by mixing, such distributions are called mixture and they arise when all (or some) parameters of a distribution vary according to some probability distribution, known by the mixing distribution. A mixture can be continuous or discrete or simply finite.

Let X be a non-negative discrete (continuous) life time of an item from a single population, with a conditional pmf (pdf) $f(x|\underline{\lambda})$ where $\underline{\lambda} = (\lambda_1, \lambda_2, \ldots \lambda_m)$ is a vector of parameters. Assume that one of the parameters λ_i has a continuous pdf $g(\theta)$ then the continuous mixture is the unconditional discrete (continuous) distribution of X which is given by

$$f(x) = \int_\theta f(x|\lambda)\, g(\theta)\, d\theta \ . \tag{1.1}$$

If the parameter λ_i has a discrete pmf $g(\theta)$, then the discrete mixture is the unconditional discrete (continuous) distribution of X which is given by

$$f(x) = \sum_\theta f(x|\lambda)\, g(\theta) \ . \tag{1.2}$$

If the item is assumed to be from one of m population then a finite mixture can be used to combine the item distributions to form a more complicated distribution. Formally, if $f_i(x)$ is the distribution of the i^{th} population, then the finite mixture is given by

$$f(x) = \sum_{i=1}^{m} p_i f_i(x) \quad \text{where} \quad 0 < p_i < 1, \ \sum_{i=1}^{m} p_i = 1 \ . \tag{1.3}$$

A well-known example of continuous-type mixture distribution is the negative binomial distribution which can be obtained as a Poisson

mixture with gamma mixing distribution; see Klugman et al. (2008) and Lemaire (1979). The negative binomial distribution provides a more flexible alternative to the Poisson distribution when the data are over-dispersed. Johnson et al. (1992) provides an inclusive survey of the applications and generalizations of the negative binomial distribution. Recently, models of mixed Poisson and mixed negative binomial distributions have been of great interest. For example, Zamani and Ismail (2010) introduced the negative binomial-Lindley which is based on mixing the distributions of negative binomial(r,q) and Lindley(θ), where the reparameterization of $q = 1 - e^{-\lambda}$ is considered. It was found that mixed distributions, in particular mixed Poisson and mixed negative binomial, provide better fit to count data compared to other distributions, for more details, one can refer to Zamani and Ismail (2010) and Lord and Geedipally (2011). Another way to deal with an over or an under-dispersed data is to fit the weighted version of the distribution instead of the original distribution itself. For instance, del Castillo and Pérez-Casany (2005) used the weighted Poisson distributions for modeling over and under-dispersed data. Weighted distributions are mainly utilized when the lifetimes do not have the same chance to be recorded. In this situation the lifetimes will follow the weighted version of the distribution instead of the original distribution itself.

Consider a non-negative random variable Y with pmf $f_Y(y)$, let w(y) be a non-negative function (weight function) with finite expectation the random variable Y_m with pmf

$$f_{Y_m}(y) = \frac{w(y)\, f_Y(y)}{E[w(y)]} \;\; ; y > 0 \; . \tag{1.4}$$

Called the weighted random variable corresponding to Y and its distribution called weighted distribution corresponding to $f_Y(y)$. The

length-biased version of a distribution is a special case of the weighted distribution when the weight function w(y)=y. A study of length-biased sampling can be found in Patil and Ord (1975). For a survey of real-life applications of length-biased distributions, see Patil and Rao (1977, 1978).

1.2 Some basic concepts in reliability

Let X be a non-negative discrete life time of a component with values in the set $N = \{0,1,2,....\}$. Thus X measures the functioning duration of the component from the beginning of its operation until its failure. Let $p_x = p(X = x)$, $x \in N$ be the pmf of X.

The *cumulative distribution function* of X, denoted by cdf, is defined by

$$F(x) = p(X \le x) = \sum_{i=0}^{x} p_i \quad . \tag{2.1}$$

Note that every cdf must satisfy four conditions:

1. $\lim_{x\to\infty} F(x) = 1$.
2. $\lim_{x\to-\infty} F(x) = 0$.
3. F(x) is non-decreasing step function.
4. F(x) is right continuous.

There are many reliability functions defined in term of the cdf, for example, the *survival function* (sf) which is the probability that the component is functioning beyond x.

$$S(x) = p(X > x) = \sum_{i=x+1}^{\infty} p_i = 1 - \sum_{i=0}^{x} p_i = 1 - F(x) \quad . \tag{2.2}$$

This quantity is sometimes called the *reliability function*.

One can note that every discrete sf must satisfy three conditions:

1. S(0) =1.
2. $\lim_{x\to\infty} S(x) = 0$.
3. S(t) is non-increasing step function.

The *conditional survival function*, $S_{X/X\succ a}(t)$, i.e. the component survives beyond x given that it has lived a is given by:

$$S_{X/X>a}(x) = \frac{P(X > x \text{ and } X > a)}{P(X > a)} = \frac{P(X > x)}{P(X > a)} = \frac{S(x)}{S(a)}, \quad x > a \ . \tag{2.3}$$

Another basic quantity in reliability is the *failure rate* (FR) which is the conditional probability of failure of the component at time x given that it did not fail before.

$$h(x) = p(X = x / X \geq x) = \frac{p(X = x)}{p(X > x-1)} = \frac{S(x-1) - S(x)}{S(x-1)} = 1 - \frac{S(x)}{S(x-1)} \ . \tag{2.4}$$

Every discrete *failure rate* must satisfy two conditions:

1. $0 \leq h(x) \leq 1$.
2. $\sum_{x=0}^{\infty} h(x) = \infty$.

The *failure rate* can take a variety of forms, for example:

1) Constant failure rate: It is as though a unit is new at each instant of time. This is known as the memory less property.
2) Increasing failure rate: The unit is subject to ageing through wear, fatigue, or accumulated damage. In practice, this is the most common case.
3) Decreasing failure rate: This is less common, but may be true, in part, of a manufacturing process where low quality components are

likely to fail early. A burn-in process may be used to remove these defective items, leaving the higher quality components which exhibit higher performance.

4) Bathtub failure rate: Here we have an initial decreasing failure rate, followed by a fairly constant period, called the useful life, and a final phase, wear out, where the failure rate increases.

In order to determine the failure rate monotonicity for continuous distribution we look at the log-concavity or the log-convexity of the distribution, see Barlow and Proschan (1981). In fact, Gupta et al. (1997) proposed the following analogous statements for discrete distributions:

1) The distribution is log-concave and therefore has an IFR if and only if $\left\{\frac{p(X=x+1)}{p(X=x)}\right\}_{x\geq 1}$ is decreasing.
2) The distribution is log-convex and therefore has a DFR if and only if $\left\{\frac{p(X=x+1)}{p(X=x)}\right\}_{x\geq 1}$ is increasing.
3) The distribution has a CFR if and only if $\left\{\frac{p(X=x+1)}{p(X=x)}\right\}_{x\geq 1}$ is constant.

The survival function can be written as a product of conditional survival probabilities

$$S(x)=\prod_{y\leq x}\frac{S(y)}{S(y-1)}=\prod_{y\leq x}[1-h(y)]\,. \tag{2.5}$$

This is the discrete case of relation for the survival function and the failure rate.

The *cumulative failure rate* is defined by

$$H(x) = \sum_{y \le x} h(y) \quad . \tag{2.6}$$

The *mean lifetime* is given by

$$\mu = E(x) = \sum_{x=0}^{\infty} x\, p_x = \sum_{x=1}^{\infty} S(x) \quad . \tag{2.7}$$

The *mean remaining life* of a component aged x, denoted by MRL, is defined to be

$$\mu(x) = E(X - x \,/\, X \ge x) = \frac{1}{S(x-1)} \sum_{y=x}^{\infty} S(y-1)\,. \tag{2.8}$$

For more details about the interrelationship between the various quantity discussed above one can refer to Lawless (2003), Barlow and Proschan (1981) and Leemis (2003).

Next, the basic concepts of discrete reliability classes will be defined and the relation between them will be discussed briefly.

Definition 1: A discrete distribution of a non-negative random variable, is an increasing (decreasing) failure rate in average, denoted by IFRA, (DFRA) if H(t)/t is non-decreasing (non-increasing) in t.

Definition 2: A discrete distribution of a non-negative random variable, is said to be new better (worse) than used, denoted by NBU (NWU) if

$$S(x+y) \le (\ge) S(x)\, S(y) \;, \qquad \forall x \ge 0,\;\; \forall y \ge 0\,. \tag{2.9}$$

The equality in (2.9) holds if and only if the distribution is the geometric.

Definition 3: A discrete distribution, of a non-negative random variable, is said to be new better (worse) than used in expectation, denoted by NBUE (NWUE) if

$$\mu(x) \le (\ge) \mu(0), \quad \forall x \ge 0. \tag{2.10}$$

The following chain implications exist among the above ageing classes:

$$\begin{array}{ccc} IFR\,(DFR) \Rightarrow IFRA\,(DFRT) \Rightarrow & & NBU\,(NWU) \\ \Downarrow & & \Downarrow \\ DMRL\,(IMRL) & \Rightarrow & NBUE(NWUE) \end{array}$$

For more details about these ageing concepts, one can refer to Abouammoh (1988), Barlow and Proschan (1981) and Bryson and Siddiqui (1969)

1.3 Estimation

Suppose that X is a random variable whose distribution is $p_x = p(X = x;\underline{\theta})$, where the form of the mass function is known except that it contains unknown parameters $\underline{\theta} = (\theta_1, \theta_2, ...\theta_m)$. Moreover, assume that the values $x_1, x_2, ...x_n$ of a random sample $X_1, X_2, ...X_n$ from $p_x = p(X = x;\underline{\theta})$ can be observed. Based on the observed sample values we estimate the value of the unknown parameters using either point or interval estimation. In point estimation we use the value of a statistic to estimate the population's parameter and this value called point estimate whereas the statistic itself called point estimator. For example, the sample mean $\bar{X}$ can be used as a point estimator of the population mean μ, in which case $\bar{x}$ is a point estimate of μ. In interval estimation we define two statistics such that their values constitute an interval for which the probability that this interval contains the unknown parameter can be determined.

1.3.1 Some properties of estimators

Numerous statistical properties of estimators can be used in order to decide which estimator is preferable. The properties that we are going to

discuss throughout this thesis are unbiasedness, lower bound of the variance of an unbiased estimator and minimum variance.

Since there are no perfect estimators that always give the exact answer, it would be reasonable that the estimator takes on average the exact value of the parameter and in this case the estimator is said to be unbiased; otherwise biased. Formally,

Definition 4: A statistic $\hat{\theta}$ is an unbiased estimator of the parameter θ if and only if $E(\hat{\theta}) = \theta$.

Since unbiased estimators are not necessarily unique we have to choose the reliable one among them, we usually take the one whose sampling distribution has the smallest variance and in this case the estimator is said to be minimum variance unbiased estimator or the best unbiased estimator.

Theorem 1: Given a random sample of size n from the distribution $f(x,\theta)$**,** if the range of x is independent of θ and $\hat{\theta}$ is an unbiased estimator of θ then under the regularity condition the following inequality holds

$$V(\hat{\theta}) \geq \frac{1}{E\left[\frac{\partial}{\partial \theta} \ln L(\theta)\right]^2} \quad , \tag{3.1}$$

where $L(\theta)$ is the likelihood function. This inequality called the Cramer-Rao lower bound of the variance of an unbiased estimator.

Remark 1: An estimator whose variance coincides with the CR lower bound is the best estimator.

Result 1: There exists an unbiased estimator $T = t(X_1, X_2,.., X_n)$ *of* $h(\theta)$ whose variance attains the CR lower bound if and only if $\frac{\partial}{\partial\theta}\ln L(\theta)$ can be expressed as

$$\frac{\partial}{\partial\theta}\ln L(\theta) = a(\theta)\left[t - h(\theta)\right] \ , \tag{3.2}$$

where $a(\theta)$ is independent of the observations.

1.3.2 Estimation methods

In this section two commonly used methods are presented briefly as well as their properties.

a) Moment method:

This method consists of equating the first few population moments to the corresponding moments of a sample.

Given a sample of size n. The k^{th} *sample moment* is given by

$$m_k = \sum_{i=1}^{n} x_i^k \Big/ n \ . \tag{3.3}$$

And the k^{th} *population moment* is given by

$$\mu_k = E(X^k) \ . \tag{3.4}$$

The method of moment estimation is based on solving the following r equations for the r parameters

$$m_k = \mu_k \quad k = 1,2,....r \ . \tag{3.5}$$

The method of moment estimators have the following properties:

1- Easy to compute compared to other methods.

2- Estimators are usually consistent.

b) Maximum likelihood method:

Given a sample of size n from a population with probability mass function $p(X=x,\underline{\theta}),\ \underline{\theta}=(\theta_1,\theta_2,.....,\theta_m)$ is a vector of unknown parameters, the *likelihood function* of this sample is defined to be

$$L(\underline{\theta})=p(X=x_1,X=x_2,....,X=x_n,\underline{\theta})=\prod_{i=1}^{n}p(X=x_i,\underline{\theta}) \ . \tag{3.6}$$

The maximum likelihood estimator of $\underline{\theta}$ is the value of $\underline{\theta}$ that maximizes $L(\underline{\theta})$.

Here, we state some important properties of the mle:

1-mle is asymptotically unbiased.

2-mle has the invariance property means if $\hat{\theta}$ is the mle of θ where θ assumed to be unidimentional then for any one to one function g the mle of $g(\theta)$ is $g(\hat{\theta})$.

For more details about this method see Mood, et al. (1985),

1.4 Goodness-of-fit techniques

This section is devoted to present goodness-of-fit methods means, techniques for examining, in particular, how well a sample of data concur with a given distribution. The formal procedure is hypotheses testing where the null hypotheses is that the given sample follows a stated distribution. In this thesis Kolmogorov-Smirnov test will be used to determine whether the proposed distributions fit the data. Goodness-of-fit techniques also involve less formal approach, in particular, graphical techniques. For example in the probability-probability plot the empirical

cumulative distribution function values are plotted against the theoretical (fitted) cdf values. In these plots a straight line suggests that the hypothesized distribution is an equitable model for the data whereas the deviation from straightening indicates that the model is inappropriate. Shapiro and Brain (1981) strongly recommended that any goodness-of-fit test should always be augmented by a probability plot.

Kolmogorov-Smirnov test

The Kolmogorov–Smirnov test is a non-parametric test which is based on the empirical distribution function.

Assume that we have a random sample of size n, $X_1, X_2,.., X_n$ drawn from a discrete distribution with a cumulative distribution function F and let $X_{(1)}, X_{(2)},.., X_{(n)}$ be the order statistics, then the *empirical cumulative distribution function*, denoted by ecdf, is defined as

$$F_n(x_{(i)}) = \frac{\#(X_j \le x_{(i)})}{n} = \frac{i}{n} \quad , \; i = 1,2...n \; , -\infty < x < \infty \, . \tag{4.1}$$

Where $\#(X_j \le x)$ is the number of the X_j's that are less than or equal to x. $F_n(x)$ represents the proportion of the sample data observations that are less than or equal to x.

If $F_o(x)$ is the hypothesized distribution function then the *K-S statistic* to test the null hypothesis $H_o : F(x) = F_o(x)$ for every x in the sample is given by

$$KS = Sup_x \left| F_n(x) - F_o(x) \right| \, . \tag{4.2}$$

The assumed distribution will be rejected if the value of the test statistic, is greater than the critical value obtained from a table.

This test has many advantages since it can be applied regardless of the assumed failure distribution of a given sample data in addition it is quite effective regardless of the sample size unlike other goodness-of-fit tests. On the other hand, it has important limitations, as only approximate result will be found if we are using discrete data. Furthermore, the distribution must be fully specified, that is if the parameters are estimated from the data then the calculated test statistics is approximately similar to the K-S statistic. For further information in this test see Law and Kelton (1991).

1.5 Scope of the thesis

The main aim of this thesis is to review most existing and commonly used discrete distributions and introduce new discrete version of the gamma distribution which is to our knowledge has not defined earlier. Moreover, this thesis introduces the generalized discrete Lindley distribution which is derived by mixing the proposed discrete gamma with the geometric. The statistical and reliability properties of the proposed distributions are investigated and estimation of the underlying parameters. Finally applications in real-life data will be given and some goodness-of-fit tests such as Kolmogorov-Smirnov test will be used to determine whether the fitted distribution fits the data applied better than other distributions such as the negative binomial distribution. the organization of the thesis will be as follows:

In this chapter we have given an overview of the general statistical and reliability concepts that need to be considered in the next chapters.

In chapter 2 a survey on discrete lifetime distributions are presented. This chapter will demonstrate that the discrete lifetime distributions can be classified into two groups, the first group consist of the discrete distributions that can be defined as the discrete counter-part of the

continuous ones. The second group consists of distributions that can be derived by using mixture of two or more discrete and continuous distributions. Their basic properties are presented and their interrelations are investigated.

Chapter 3 introduces new discrete distribution which can be viewed as the discrete version of the gamma distribution. The statistical and reliability properties of the discrete gamma distribution are established. Its relationships with the geometric and negative binomial distributions are explored. Furthermore, the size-biased version of this distribution is derived. Finally, some mixtures of the discrete gamma are derived and studied.

In Chapter 4 two estimation methods namely, the maximum likelihood method and method of moment are used to estimate the parameters of the discrete gamma and the generalized discrete Lindley distributions. Applications in real-life data are also considered. In addition, some heuristic and analytical tests to determine the goodness-of-fit will be discussed. In particular, K-S test augmented by a probability plot.

Chapter 2

A Survey on discrete lifetime distributions

This chapter presents an inclusive survey of discrete lifetime distributions. Here, the main statistical and reliability properties of each model are derived. This chapter shows that the discrete lifetime distributions can be classified into different groups. One group consists of discrete distributions that can be defined as discrete counter-parts of continuous ones based on time discretization. Other group consists of the discrete distributions defined without any continuous counter-part. Also, some of discrete distributions can be derived based on mixing. Moreover the length-biased version of some discrete distributions is briefly discussed.

2.1 Discrete distributions with continuous counter-parts

This can be done either by considering the characteristic property of the continuous distribution and building the similar one in discrete time, such as the geometric distribution. Alternatively, by considering the integer part of the continuous time as the discrete time.

2.1.1 Geometric distribution

Definition 1: A discrete non-negative random variable X, is said to have Geometric distribution with parameter q, denoted by G(q), if its pmf is given by

$$p_x = p(X = x) = p\,q^x, \quad x = 0, 1, 2, \ldots, \quad 0 < q < 1, \quad p = 1 - q. \tag{1.1}$$

The mean and variance of the G(q) are given by:

$$E(X)=\frac{q}{p} \tag{1.2}$$

$$V(X)=\frac{q}{p^2} \quad . \tag{1.3}$$

One can note that the variance is greater than the mean therefore the G is an over-dispersed distribution.

The cdf, sf, FR and MRL of the G(q) are given by:

$$F(x)=1-q^{x+1} \tag{1.4}$$

$$S(x)=q^{x+1} \tag{1.5}$$

$$h(x)=p \tag{1.6}$$

$$\mu(x)=\frac{1}{p} \quad , \tag{1.7}$$

respectively. This distribution is considered to be the discrete counter-part of the exponential distribution since it has the memory less property which is given by

$$p(X>x+y \setminus X>x)=p(X>y) \quad \forall x,y\in\{0,1,2,.....\} \, . \tag{1.8}$$

Also it has a constant failure rate and it is unimodal with zero mode.

Using (1.1) and (1.2) the length-biased geometric is given by

$$p_{o_x}=\frac{x\,p_x}{E(X)}=p^2xq^{x-1} \quad ,x=1,2,... \quad . \tag{1.9}$$

2.1.2 Discrete uniform distribution

Definition 2: A discrete random variable X is defined to have a discrete uniform distribution with parameter N, if its pmf is given by

$$p_x = p(X = x) = \frac{1}{N}, \quad x = 1, 2, ..., N \ , \quad N = 1, 2.... \quad . \tag{1.10}$$

The mean and variance of this distribution are given by:

$$E(X) = \frac{N+1}{2} \tag{1.11}$$

$$V(X) = \frac{(N+1)(N-1)}{12} \quad . \tag{1.12}$$

The cdf, sf and FR and MRL of this distribution are given by:

$$F(x) = \frac{x}{N} \tag{1.13}$$

$$S(x) = \frac{N-x}{N} \tag{1.14}$$

$$h(x) = \frac{1}{N-x+1} \tag{1.15}$$

$$\mu(x) = \frac{2+N-x}{2} \quad . \tag{1.16}$$

This distribution has an IFR for any $x \in [1, N]$ and any positive integer N and therefore, it has a DMRL.

2.1.3 Negative binomial distribution

The negative binomial distribution is the sum of r iid geometric random variables. Based on this property the negative binomial distribution is considered by many researchers, to be the discrete counter-part of the gamma distribution.

Definition 3: A discrete random variable X, is said to have a negative binomial distribution with parameters (r,q), denoted by NB(r,q), if its pmf is given by

$$p_x = \binom{r+x-1}{x} q^x p^r \quad ; 0 < q < 1 , \ \ p = 1 - q , \ r = 1, 2, ... \ \ , \ x = 0, 1, 2, ... \tag{1.17}$$

The mean and the variance of the NB(r,q) are given by:

$$E(X) = \frac{rq}{p} \tag{1.18}$$

$$V(X) = \frac{rq}{p^2} \ . \tag{1.19}$$

One can note that, the variance is greater than the mean. Therefore, the NB can be a good alternative to distributions that can't deal with an over-dispersed data.

To compute the mode, we note that, $p_x > p_{x-1}$ if and only if $x < \frac{(r-1)q}{p}$.

Therefore, the mode of the NB is equal to the integer part of $\frac{(r-1)q}{p}$.

The sf and FR of the NB(r,q) can not be written in a compact form. In order to determine the failure rate monotonicity, we are going to look into the log-concavity of the NB(r,q). Now, consider the function

$$b(x,r) = \frac{p(X = x+1)}{p(X = x)} = q\frac{r+x}{x+1} \ . \tag{1.20}$$

Its derivative is given by

$$\frac{db(x,r)}{dx} = \frac{q(1-r)}{(x+1)^2} < 0 \qquad for\, r \geq 2.$$

Note that $b(x,r)$ is a decreasing function in x for $x \geq 1, r \geq 2$ thus the NB(r,q) is log-concave. Since the expression of the failure rate of the NB (r,q) is not given in a compact form, one can use the log-concavity property to state that h(x) is an increasing failure rate (IFR), for $r \geq 2$. For r=1 the distribution reduced to the G(q) and hence the failure rate is constant.

Using (1.17) and (1.18) the length-biased NB is given by

$$p_{o_x} = \frac{x\, p_x}{E(X)} = \binom{r+x-1}{x-1} q^{x-1}\, p^{r+1} \;, x = 1,2,\ldots \;. \tag{1.21}$$

For the properties of the length-biased NB one can refer to Mir (2009).

2.1.4 Type I discrete Weibull

This model was introduced by Nakagawa-Osaki (1975). It is obtained from the continuous Weibull by time discretization based on the similarity of expressions of the survival function between the discrete and continuous time.

Definition 4: A discrete random variable X, is said to have a type I discrete Weibull distribution with parameters (α,q), denoted by WI(α,q), if its pmf is given by

$$p_x = q^{x^\alpha} - q^{(x+1)^\alpha} \qquad ; \; 0 < q < 1 \;,\; \alpha > 0 \;,\; x = 0,1,2,\ldots. \tag{1.22}$$

Here α is the shape parameter and q is the probability of surviving the first demand.

The sf and FR of the WI(α,q) are given by

$$S(x) = q^{(x+1)^\alpha} \tag{1.23}$$

$$h(x) = 1 - q^{(x+1)^{\alpha} - x^{\alpha}} . \tag{1.24}$$

If α=1 then the WI(α,q) reduced to the G(q) and the failure rate is constant. Similarly if $\alpha > 1$, the failure rate is increasing and if $0 < \alpha < 1$ then the failure rate is decreasing.

This distribution is considered to be the best probabilistic analogies with the continuous Weibull since it is simple, flexible according to the selection of α and its parameter have a physical meaning.

2.1.5 Type II discrete Weibull distribution

This model was introduced by Stein and Dattero (1984) and it is based on the similarity of expressions of the failure rates for discrete and continuous Weibull.

Definition 5: A discrete random variable X, is said to have a type II discrete Weibull distribution with parameters (α,m), denoted by WII(α,m), if its pmf is given by

$$p_x = \frac{x^{\alpha-1}}{m} \prod_{i=1}^{x-1} \left[1 - \frac{i^{\alpha-1}}{m} \right] \quad ; \; m = 1,2,... \;\; , \;\; \alpha > 0 \;\; , \;\; x = 1,2,....,m . \tag{1.25}$$

Here α is the shape parameter and m is the maximal lifetime of the system.

The sf and FR of the WII(α,m) are given by

$$S(x) = \prod_{i=1}^{\inf(x,m)} \left[1 - \frac{i}{m}^{\alpha-1} \right] \tag{1.26}$$

$$h(x) = \frac{x}{m}^{\alpha-1} . \tag{1.27}$$

The distribution has an IFR for $\alpha > 1$, DFR for $0 < \alpha < 1$ and CFR for α=1

This distribution is not very useful practically because it has a bounded support which is not always realistic.

2.1.6 Type III discrete Weibull distribution

This model was defined by Padgett and Spurrier (1985).

Definition 6: A discrete random variable X, is said to have a type III discrete Weibull distribution with parameters (α,c), denoted by WIII(α,c), if its pmf is given by

$$p_x = e^{-c\sum_{i=1}^{x} i^{\alpha}} \left[1 - e^{-c(x+1)^{\alpha}}\right] \quad ; \alpha \in R \ , \ c > 0 \ , \ x = 0,1,2,.... \ . \tag{1.28}$$

Here α is the shape parameter and c is related to the probability of failure at the first demand since $p_o = 1 - e^{-c}$.

The sf and FR of the WIII(α,c) are given by

$$S(x) = e^{-c\sum_{j=1}^{x+1} j^{\alpha}} \tag{1.29}$$

$$h(x) = 1 - e^{-c(x+1)^{\alpha}} \ . \tag{1.30}$$

If α=0 the WIII(α,c) reduced to the $G(e^{-c})$ and the failure rate is constant. Similarly if $\alpha > 1$ then the distribution has an IFR and if $\alpha < 1$ then it has a DFR.

2.2 Discrete distributions without continuous counter-part

This section presents some discrete distributions that are either independently defined or derived based on mixing.

2.2.1 Binomial distribution

Typically, a binomial random variable is the number of successes in a series of trials.

Definition 7: A discrete random variable X is said to follow a Binomial distribution with parameters n and p, if its pmf is given by

$$p_x = \binom{n}{x} p^x q^{n-x}, \; x = 0,1,2,\ldots,n \tag{2.1}$$

here, $n = 1,2,\ldots.$, $0 < p < 1$, $p = 1 - q$.

The mean and the variance of this distribution are given by:

$$E(X) = np \tag{2.2}$$

$$V(X) = npq \; . \tag{2.3}$$

One can note that the variance is less than the mean therefore the binomial is an over-dispersed distribution.

To compute the mode, we note that, $p_x > p_{x-1}$ if and only if $x < (n+1)p$.

Therefore, the mode of the binomial distribution is equal to the integer part of $(n+1)p$. But, if $(n+1)p$ is an integer then the distribution has two consecutive modes, $(n+1)p$ and $(n+1)p-1$.

In order to determine the failure rate monotonicity we are going to look into the log-concavity of this distribution. Now, consider the function

$$b(x,n) = \frac{p(X = x+1)}{p(X = x)} = \frac{n-x}{x+1}. \tag{2.4}$$

Its derivative is given by

$$\frac{db(x,n)}{dx} = \frac{-(n+1)}{(x+1)^2} < 0 \qquad for\; n \geq 1 \; .$$

Note that $b(x,n)$ is a decreasing function in x for $x \geq 0, n \geq 1$ thus the binomial distribution is log-concave and hence, it has an IFR.

Using (2.1) and (2.2), the length-biased Binomial distribution is given by

$$p_{o_x} = \frac{x\, p_x}{E(X)} = \binom{n-1}{x-1} p^{x-1} q^{n-x}, x = 1,2,\ldots \; . \tag{2.5}$$

2.2.2 Poisson distribution

The Poisson distribution is used to model the number of events occurring within a given time interval, such as the number of telephone calls at a business center in a specific time period.

Definition 8: A discrete random variable X is said to follow a Poisson distribution with parameter λ, if its pmf is given by

$$p_x = \frac{e^{-\lambda}\lambda^x}{x!} \;,\; x = 0,1,2,\dots \;,\; \lambda > 0 \;. \tag{2.6}$$

Here, λ is the shape parameter which indicates the average number of events in the given time interval.

To compute the mode, we note that, $p_x > p_{x-1}$ if and only if $x < \lambda$. Therefore, the mode of the Poisson distribution is equal to the integer part of λ. But if λ is an integer then the distribution has two consecutive modes, λ and λ-1.

It is well known that the mean and the variance of the Poisson distribution are both equal to λ. Therefore, it is constrained by its equi-dispersion property, which makes it less than ideal for modeling count data that often exhibit over or under-dispersion. If the data is under-dispersed, then neither the Poisson nor the negative binomial distributions provide adequate approximations. Several distributions have been proposed for modeling under-dispersion. For instance, the length-biased version of the Poisson which is given by

$$p(X = x) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \;;\; \lambda > 0 \;,\; x = 1,2,\dots. \tag{2.7}$$

This distribution is an under-dispersed distribution since its mean $E(X) = 1+\lambda$ is bigger than its variance $V(X) = \lambda$.

2.2.3 Poisson-Lindley distribution

The Poisson-Lindley distribution was introduced by Sankaran (1970) to model count data. It is derived from the Poisson distribution when its parameter λ follows a Lindley (1958) distribution with pdf

$$f(\lambda)=\frac{\theta^2}{\theta+1}(1+\lambda)e^{-\theta\lambda} \;\; ; \;\; \theta>0 \;, \lambda>0 \;. \tag{2.8}$$

The mgf of the Lindley distribution is given as

$$M_{\lambda}(t)=\frac{\theta^2}{\theta+1}\frac{\theta-t+1}{(\theta-t)^2} \;. \tag{2.9}$$

Definition 9: A discrete random variable X, is said to have a Poisson-Lindley distribution with parameter θ, denoted by PL(θ), if it satisfies the stochastic representation:

$$X/\lambda \sim Poisson(\lambda) \;\; and \;\; \lambda \sim Lindley(\theta) \;\; .$$

The conditional distribution of $X \mid \lambda$ is given by

$$p(X=x \mid \lambda)=\frac{e^{-\lambda}\lambda^x}{x!} \;\; , x=0,1,2,\ldots, \;\; \lambda>0 \;\; . \tag{2.10}$$

Now, using (2.8) and (2.10), the form of the mixing that leads to the PL(θ) is given by

$$\begin{aligned} p(X=x) &= \int_0^{\infty} p(X=x \mid \lambda) f(\lambda)\, d\lambda \\ &= \int_0^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}\frac{\theta^2}{\theta+1}(1+\lambda)e^{-\lambda\theta}\, d\lambda \\ &= \frac{\theta^2}{(\theta+1)x!}\int_0^{\infty} e^{-\lambda(\theta+1)}\lambda^x (1+\lambda)\, d\lambda \\ &= \frac{\theta^2}{\theta+1}\left[\frac{1}{(\theta+1)^{x+1}}+\frac{x+1}{(\theta+1)^{x+2}}\right] \\ &= \frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}} \,. \end{aligned}$$

Hence, the pmf of the PL(θ) is given by

$$p(X=x)=\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}} \quad ; \quad \theta>0 \;\; , \; x=0,1,2,.... \; . \tag{2.11}$$

The mean and variance of the PL(θ) are given by

$$\mu=\frac{\theta+2}{\theta(\theta+1)} \tag{2.12}$$

$$\sigma^2=\frac{2+6\theta+4\theta^2+\theta^3}{\theta^2(\theta+1)^2}\,. \tag{2.13}$$

One can note that, the PL(θ) is an over-dispersed distribution since $V(X)-E(X)=\frac{2+4\theta+\theta^2}{\theta^2(1+\theta)^2}>0$ for any $\theta>0$.

It has been shown that in many ways, Lindley is a better distribution compared to the exponential. Hence, it is expected that the PL(θ) provides a better fit compared to the Poisson-exponential. For more details one can refer to Ghitany et al. (2008).

Now suppose that the lifetimes of a given sample of items follow a PL(θ) distribution and the items are selected according to their life lengths means they are not equally likely to be selected then these lifetimes will have a length-biased Poisson–Lindley distribution instead of their original distribution. Using (2.11) and (2.12) the length-biased Poisson-Lindley distribution, denoted by LBPL(θ), is given by

$$p_o(X=x)=\frac{x\,p(X=x)}{\mu}=\frac{\theta^3}{\theta+2}\,\frac{x(x+\theta+2)}{(\theta+1)^{x+2}} \quad . \tag{2.14}$$

The LBPL distribution also arises as a mixture of the length-biased Poisson distribution given in (2.7) where the mixing distribution of λ is the length-biased Lindley distribution with pdf

$$f_o(\lambda)=\frac{\theta^3}{\theta+2}\lambda(1+\lambda)e^{-\theta\lambda}; \quad \lambda>0, \quad \theta>0 \; .$$

The mean and the variance of the LBPL(θ), are given by:

$$E(X)=\frac{\theta^2+4\theta+6}{\theta(\theta+2)} \tag{2.15}$$

$$V(X)=\frac{2(\theta^3+6\theta^2+12\theta+6)}{\theta^2(\theta+2)^2}. \tag{2.16}$$

One can note that,

$V(X)-E(X)=-\frac{\theta^4+4\theta^3+2\theta^2-12\theta-12}{\theta^2(\theta+2)^2}=(>)(<)0$ if and only if $\theta=(>)(<)1.6711$. This means that the LBPL(θ) is equi-dispersed, over-dispersed and under-dispersed for $\theta=(>)(<)1.6711$.

Next, we study the log-concavity of the LBPL(θ), let's consider the function

$$b(x,\theta)=\frac{p(X=x+1)}{p(X=x)}=\frac{1}{\theta+1}(1+\frac{1}{x})(1+\frac{1}{x+\theta+2}) \tag{2.17}$$

Note that $b(x,\theta)$ is a decreasing function in x thus, the LBPL(θ) is log-concave. Therefore it is unimodal, has an IFR and a DMRL. For more details about this distribution, one can refer to Ghitany and Almutairi (2008).

2.2.4 Negative binomial-Lindley distribution

The negative binomial-Lindley distribution was introduced by Zamani and Ismail (2010) to model count data. The distribution is based on mixing the distributions of NB(r,q) and Lindley(θ), where the reparameterization of p = exp(-λ) is considered.

Definition 10: A discrete random variable X, is said to have a negative binomial-Lindley distribution with parameters(r,θ), denoted by NBL(r,θ), if it satisfies the stochastic representation

$$X/\lambda \sim NB(r,1-e^{-\lambda}) \ \ and \ \ \lambda \sim L(\theta) \quad where\ ,\ r>0 \ and\ \theta>0\,.$$

Using (1.17) the conditional distribution of $X\,|\,\lambda$ is given by

$$p(X=x\,|\,\lambda)=\binom{r+x-1}{x}(1-e^{-\lambda})^x e^{-\lambda r} \quad ; \quad r=1,2,... \ \ , \ x=0,1,2,.... \,. \tag{2.18}$$

By using the binomial theorem we get

$$p(X=x\,|\,\lambda)=\binom{r+x-1}{x}e^{-\lambda r}\sum_{j=0}^{x}\binom{x}{j}(-1)^j e^{-\lambda j} \quad . \tag{2.19}$$

Now, using (2.19) and (2.8), The form of the mixing that leads to the NBL(r,θ) is given by

$$\begin{aligned} p(X=x) &= \int_0^\infty p(X=x\,|\,\lambda)\, f(\lambda)\, d\lambda \\ &= \binom{r+x-1}{x}\sum_{j=0}^{x}\left[\binom{x}{j}(-1)^j \int_0^\infty e^{-\lambda(r+j)} f(\lambda)d\lambda\right] \\ &= \binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j M_\lambda(-r-j)\,. \end{aligned}$$

Finally, using the mgf of the Lindley distribution given in (2.9), leads to the pmf of the NBL(r,θ) which is given by

$$p_x=\frac{\theta^2}{\theta+1}\binom{r+x-1}{x}\sum_{j=0}^{k}\binom{k}{j}(-1)^j\frac{\theta+r+j+1}{(\theta+r+j)^2} \qquad , x=0,1,2,... \ . \tag{2.20}$$

Next, we are going to study the log-concavity of the NBL(θ), we consider the function

$$b(x,r)=\frac{p(X=x+1)}{p(X=x)}=\frac{r+x}{x+1}\,. \tag{2.21}$$

Its derivative is given by

$$\frac{db(x,r)}{dx}=\frac{1-r}{(x+1)^2}<0 \qquad for\ r\geq 2\,.$$

Note that $b(x,r)$ is a decreasing function in x for $x\geq 1, r\geq 2$ thus the NBL is log-concave. Therefore it is unimodal, hence it has an IFR and a DMRL.

Based on the results given by Zamani (2010), it is shown that the NB-L distribution can be considered as an alternative for modelling count data

since it provides a better fit compared to the Poisson and the negative binomial distributions.

2.3 Interrelationships between discrete distributions

Here, the main interrelationships between various distributions reviewed earlier are pointed out in the following:

a) In fact the relationships between the geometric and other discrete distribution can be stated as:

   1. The NB(1,q) is the G(q).

   2. The WI(1,q) is the G(q).

   3. The WII(1,m) is the shifted geometric distribution G(1-1/m).

   4. The WIII(0,c) is the $G(e^{-c})$.

b) The negative binomial distribution appears to be a continuous mixture of Poisson distributions where the mixing distribution of the Poisson rate is a gamma distribution. Formally, the following stochastic representation must be satisfied

$$X/\lambda \sim Poisson(\lambda) \;\; and \;\; \lambda \sim gamma(r, \frac{q}{p}) \quad .$$

The form of the mixing that leads to the NB(r,q), is given by

$$
\begin{aligned}
p(X=x) &= \int_0^\infty p(X=x\,|\,\lambda)\, f(\lambda)\, d\lambda \\
&= \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\,\frac{p^r}{q^r\Gamma(r)}\,\lambda^{r-1} e^{-\frac{\lambda p}{q}}\, d\lambda \\
&= \frac{p^r q^{-r}}{x!\Gamma(r)} \int_0^\infty \lambda^{r+x-1} e^{-\frac{\lambda}{q}}\, d\lambda \\
&= \frac{p^r q^{-r}}{x!\Gamma(r)}\, q^{r+x}\Gamma(r+x) \\
&= \frac{\Gamma(r+x)}{x!\Gamma(r)}\, q^x p^r \; .
\end{aligned}
$$

For more details about the NB distribution one can refer to the comprehensive survey of the applications and generalizations of the negative binomial distributions given by Johnson et al. (1992)

c) The negative binomial distribution converges to the Poisson distribution where the reparameterization of $q = \frac{\lambda}{r+\lambda}$ is considered. Formally,

$$
\begin{aligned}
\lim_{r\to\infty} NB(r,q) &= \lim_{r\to\infty}\left[\frac{\Gamma(r+x)}{(\lambda+r)^x\,\Gamma(r)}\,\frac{\lambda^x}{x!}\,\frac{1}{\left(1+\frac{\lambda}{r}\right)^r}\right] \\
&= 1.\frac{\lambda^x}{x!}\,\frac{1}{e^{\lambda}} \\
&= Poisson(\lambda).
\end{aligned}
$$

d) The normal distribution can be used to approximate the binomial distribution with the presence of correction for continuity adjustment. Since the skeweness of the binomial is given by $\frac{1-2p}{\sqrt{npq}}$ therefore, the binomial distribution is symmetric like the normal distribution whenever $p = 0.5$. However, the closer p is to 0.5 and the larger

sample size n, the more symmetric the distribution becomes. As a general rule the approximation can be used whenever np and nq are at least 5. The normal approximation to the binomial is given by

$$Z = \frac{X_a - np}{\sqrt{npq}},$$

where X_a is the adjusted number of successes for the binomial random variable X, such that $X_a = X \pm 0.5$, as appropriate.

Similarly, The normal distribution can also be used to approximate the Poisson distribution whenever the parameter $\lambda \geq 5$. The normal approximation to the Poisson is given by

$$Z = \frac{X_a - \lambda}{\sqrt{\lambda}},$$

where $X_a = X \pm 0.5$, as appropriate.

Chapter 3

Discrete Gamma Distribution

In this chapter we introduce a new discrete distribution, namely the discrete gamma distribution. It is defined to be the discrete counter-part of the continuous gamma. In fact, many authors have accepted the negative binomial distribution as an approximation or has similarity with the expected discrete gamma. For more details, one can see Johnson et al. (1992). Here the main statistical and reliability properties of the new distribution are obtained. In addition, the relationships of this distribution with the geometric and the negative binomial are investigated. Moreover, some mixtures of the discrete gamma are derived and studied.

3.1 Basic definition

The pdf of the usual continuous gamma is given by

$$f(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \; ; \; x \ge 0 \, , \; \alpha > 0 \, , \; \beta > 0 \; .$$

This distribution is shown by many authors to play a significant role in statistical modeling and applied probability. Examples of events that might be modeled by the gamma distribution include the amount of rainfall accumulated in a reservoir, the size of loan defaults or aggregate insurance claims and the load on web servers.

The discrete gamma distribution is defined based on discretizing x and α and considering the substitution $q = e^{-\beta}$. Note that, α in the $DGD(\alpha, q)$ can be any positive number but we will study the case when α is a positive integer.

Definition 1: A discrete non-negative random variable X, is said to have a discrete gamma distribution with parameters (α,q), denoted by $DGD(\alpha,q)$, if its pmf is given by

$$p(\alpha,q) = p(X=x) = C\,p^{\alpha}\,x^{\alpha-1}\,q^{x}\,;\; x=0,1,2\ldots,\;\; \alpha=1,2\ldots\ldots,\;\; 0<q<1 \qquad (1.1)$$

where C=1 if α=1 and $C=\dfrac{1}{q\sum_{i=0}^{\alpha-2}A(\alpha-1,i)\,q^{i}}$, for $\alpha \geq 2$.

Here, $p=1-q$ and $A(n,m)$ is known by the Eulerian number and it is given by

$$A(n,m) = \sum_{k=0}^{m}(-1)^{k}\binom{n+1}{k}(m+1-k)^{n} \quad .$$

For properties of the Eulerian number see Hirzebruch (2008).

Remark 1: The pmf of the $DGD(\alpha,q)$ with α=1 is the G(q).

The following are special cases of the $DGD(\alpha,q)$ when α=1,2,3 and 4:

$$p(1,q) = p\,q^{x} \qquad (1.2)$$

$$p(2,q) = \frac{1}{q}\,x\,p^{2}\,q^{x} \qquad (1.3)$$

$$p(3,q) = \frac{1}{q(1+q)}\,x^{2}\,p^{3}\,q^{x} \qquad (1.4)$$

$$p(4,q)=\frac{1}{q(1+4q+q^2)}x^3\,p^4\,q^x \quad , \tag{1.5}$$

respectively. In order to give a clear idea about the shape of the $DGD(\alpha,q)$, it is reasonable to plot the pmf for different values of α and q. Figures (1), (2) and (3) show the pmf of the $DGD(\alpha,q)$.

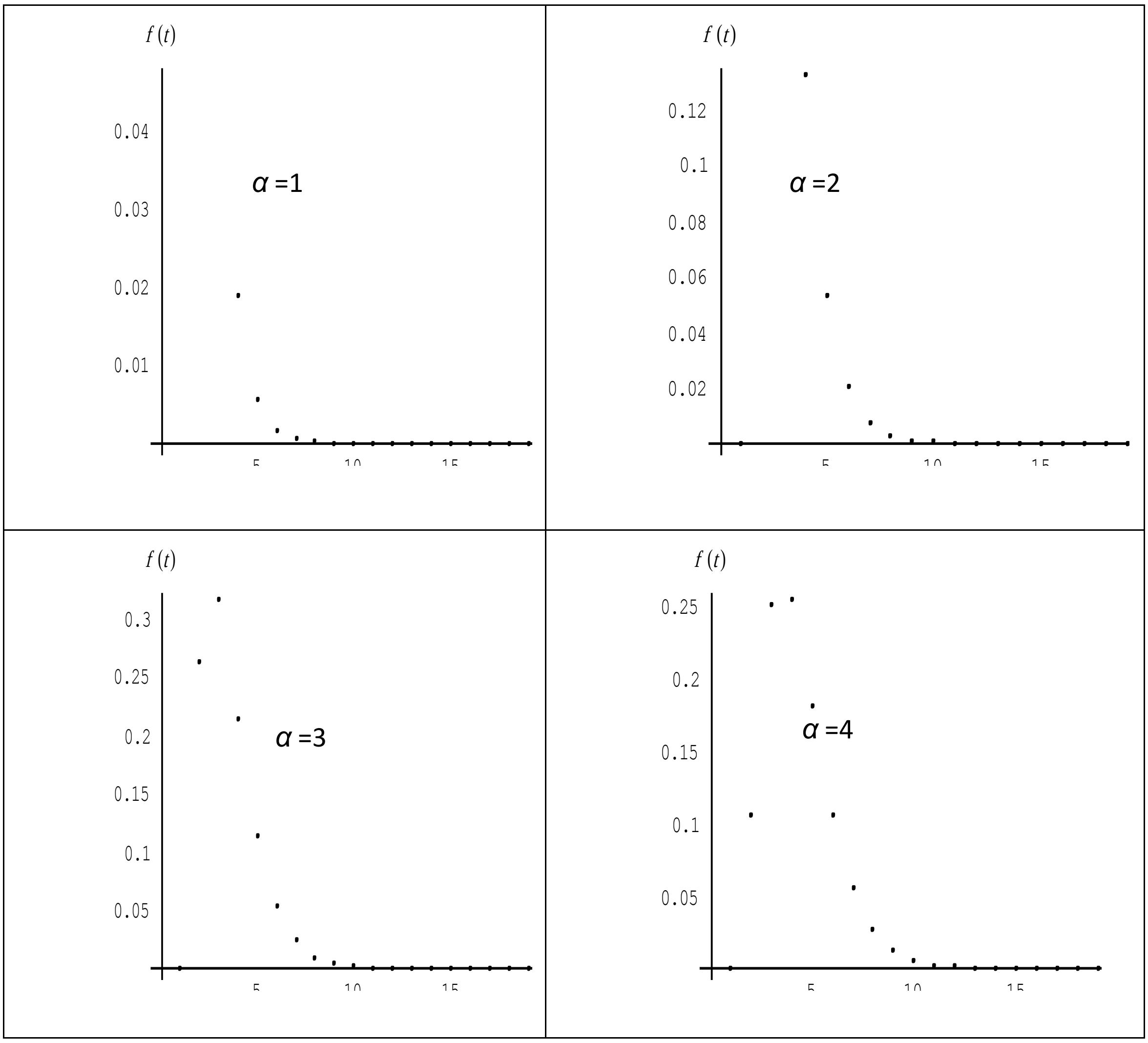


Figure 1.pmf of DGD(α ,0.3)

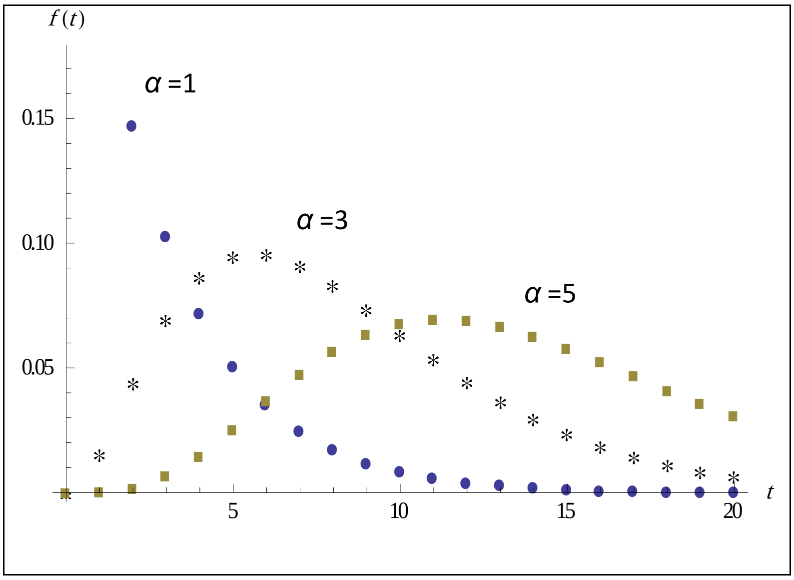


Figure 2.pmf of DGD(α ,0.7)

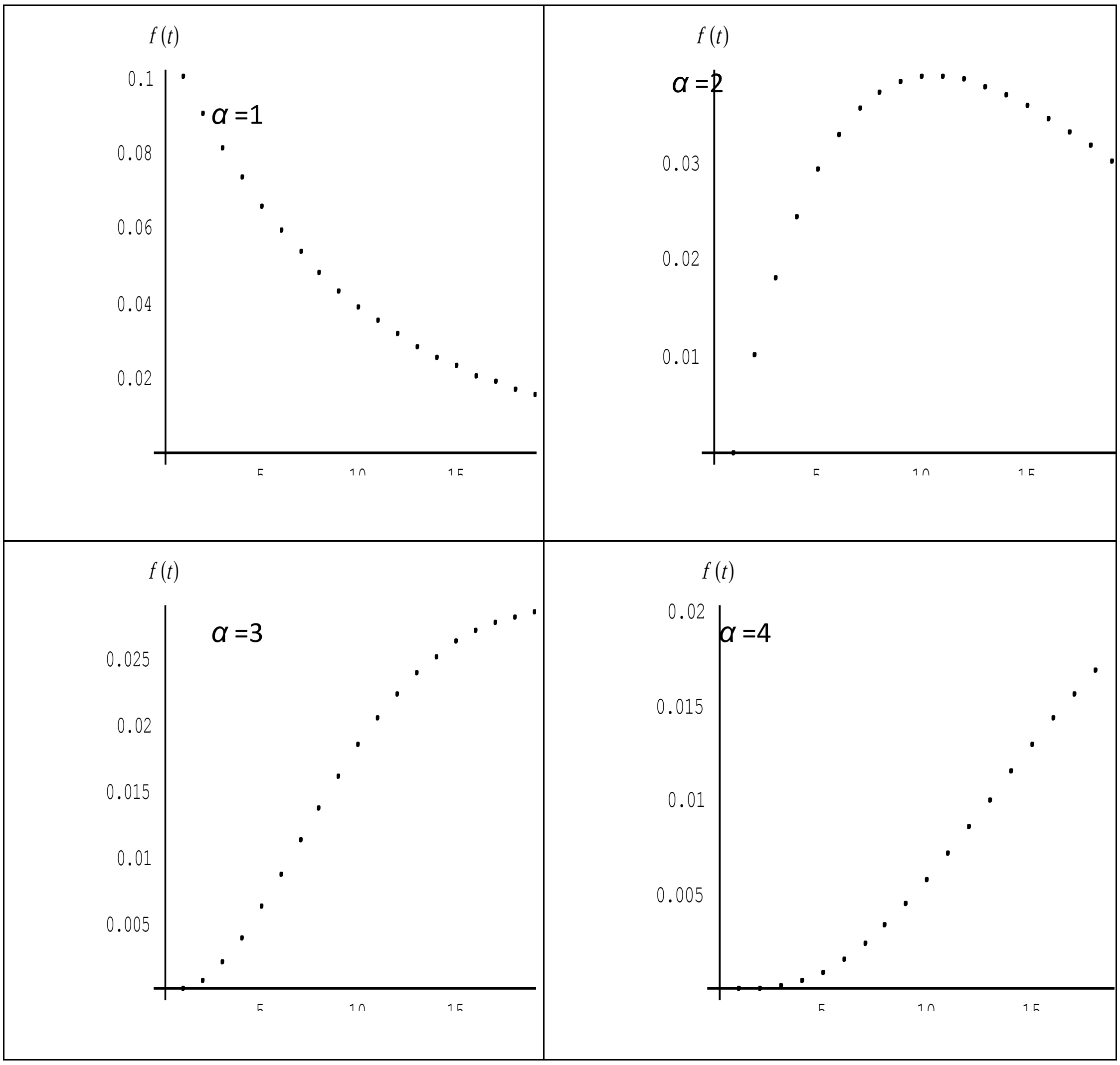


Figure 3. pmf of DGD(α ,0.9)

From the above figures, it is clear that the $DGD(\alpha, q)$ is right skewed and its skewness decreases as α increases. Also one can note that the peak of the $DGD(\alpha, q)$ becomes less sharper while its tail becomes shorter as α increases. This means that the kurtosis decreases as α increases.

3.2 Statistical properties of the DGD(α,q)

In this section the main statistical properties of the $DGD(\alpha, q)$ are obtained.

(i) The k^{th} moment of the $DGD(\alpha, q)$ is found to be:

$$E(X^k) = \frac{\sum_{i=0}^{\alpha+k-2} A(\alpha+k-1, i) q^i}{p^k \sum_{i=0}^{\alpha-2} A(\alpha-1, i) q^i}. \tag{2.1}$$

(ii)The mean and the second moment are given by:

$$\mu = E(X) = \frac{\sum_{i=0}^{\alpha-1} A(\alpha, i) q^i}{p \sum_{i=0}^{\alpha-2} A(\alpha-1, i) q^i} \tag{2.2}$$

$$E(X^2) = \frac{\sum_{i=0}^{\alpha} A(\alpha+1, i) q^i}{p^2 \sum_{i=0}^{\alpha-2} A(\alpha-1, i) q^i}, \tag{2.3}$$

respectively.

(iii) The variance is evaluated based on (2.2) and (2.3) by

$$V(X) = E(X^2) - [E(X)]^2. \tag{2.4}$$

Next, we present the mean, the second moment and the variance of the DGD(2,q)

$$E(X) = \frac{q+1}{p} \tag{2.5}$$

$$E(X^2)=\frac{1+4q+q^2}{p^2} \tag{2.6}$$

$$V(X)=\frac{2q}{p^2} \quad , \tag{2.7}$$

respectively.

Whereas, for the DGD(3,q) the mean, the second moment and the variance are given by

$$E(X)=\frac{1+4q+q^2}{(1-q)(1+q)} \tag{2.8}$$

$$E(X^2)=\frac{1+11q+11q^2+q^3}{(1+q)(1-q)^2} \tag{2.9}$$

$$V(X)=\frac{4q(1+q+q^2)}{(1-q^2)^2} \quad , \tag{2.10}$$

respectively.

Also, the moment generating function (mgf) of the $DGD(\alpha,q)$ is found to be:

$$E(e^{tX})=\frac{e^t(1-q)^\alpha[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)(qe^t)^i]}{(1-qe^t)^\alpha[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i]} \ . \tag{2.11}$$

Hence, the factorial mgf of the $DGD(\alpha,q)$ is given by:

$$E(s^X)=\frac{s(1-q)^\alpha[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)(qs)^i]}{(1-qs)^\alpha[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i]} \ . \tag{2.12}$$

Remark 2: Unlike the Poisson and the NB(r,q), the $DGD(\alpha, q)$ can handle both over and under-dispersed data.

Using (2.5) and (2.7)

$V(X)-E(X)=\dfrac{q^2+2q-1}{p^2}=(>)(<)0$ if and only if $q=(>)(<)0.4142$.

This means that the $DGD(2, q)$ is equi-dispersed, over-dispersed and under-dispersed for $q=(>)(<)0.4142$, respectively.

Similarly, using (2.8) and (2.10)

$V(X)-E(X)=\dfrac{q^4+8q^3+4q^2-1}{(1-q^2)^2}=(>)(<)0$ if and only if $q=(>)(<)0.3744$.

This means that the $DGD(3, q)$ is equi-dispersed, over-dispersed and under-dispersed for $q=(>)(<)0.3744$, respectively.

3.3 Reliability Properties of the DGD(α,q)

In this section the basic reliability properties of the $DGD(\alpha, q)$ are investigated.

The cdf, sf, FR and MRL of the $DGD(2, q)$ are given by:

$$F(x)=1-q^x-xq^x+xq^{x+1} \tag{3.1}$$

$$S(x)=q^x[(1+x)-xq] \tag{3.2}$$

$$h(x)=\frac{x(1-q)^2}{x-(x-1)q} \tag{3.3}$$

$$\mu(x) = \frac{x-(x-2)q}{(1-q)[x-(x-1)q]} \quad , \tag{3.4}$$

respectively.

The cdf, sf, FR and MRL of the $DGD(3,q)$ are given by:

$$F(x) = \frac{1+q-(1+x)^2 q^x - (1-2x-2x^2)q^{x+1} - x^2 q^{x+2}}{(1+q)} \tag{3.5}$$

$$S(x) = \frac{q^x[(1+x)^2 + (1-2x-2x^2)q + x^2 q^2]}{(1+q)} \tag{3.6}$$

$$h(x) = \frac{x^2(1-q)^3}{x^2 + (1+2x-2x^2)q + (x-1)^2 q^2} \tag{3.7}$$

$$\mu(x) = \frac{x^2 - 2(x^2-2x-1)q + (x-2)^2 q^2}{(1-q)[x^2 + (1+2x-2x^2)q + (x-1)^2 q^2]} \quad , \tag{3.8}$$

respectively.

The next result shows the failure rate behavior of the $DGD(\alpha,q)$ for $\alpha \geq 2$.

Theorem 1: If $X \sim DGD(\alpha,q)$, then X has an IFR for $x \geq 1, \alpha \geq 2$.

Proof: Since the expression of the failure rate of $DGD(\alpha,q)$ for any α is not given in a compact form, we are going to look into the log- concavity of the $DGD(\alpha,q)$ in order to determine the FR monotonicity. Now, consider the function

$$b(x,\alpha)=\frac{p(X=x+1)}{p(X=x)}=q\left(\frac{x+1}{x}\right)^{\alpha-1}.$$

Its derivative is given by

$$\frac{db(x,\alpha)}{dx}=-\frac{\alpha-1}{x^2}\left(\frac{x+1}{x}\right)^{\alpha-2}<0 \qquad for\ \alpha\geq 2$$

Note that $b(x,\alpha)$ is a decreasing function in x for $x\geq 1, \alpha\geq 2$ thus, the $DGD(\alpha,q)$ is log-concave and hence the FR is increasing.

Here, we have not tackled the case of $\alpha<1$ in spite of the fact that for $\alpha<1$ the $DGD(\alpha,q)$ will have a decreasing failure rate, since $b'(x,\alpha)$ is positive. This means that the $DGD(\alpha,q)$ has the same aging property of the continuous gamma for different values of the shape parameter α. For more details about discrete failure rate monotonic properties, see Gupta et al. (1997).

Remark 3: It is a general fact that if the distribution has an IFR then it has a DMRL. Which means in the case of the $DGD(\alpha,q)$ the mean remaining life is a decreasing function for $x\geq 1, \alpha\geq 2$. Furthermore it has an IFRA as well as NBU and hence NBUE properties.

Figures (4) and (5) represent the graphs of the FR for different values of α when q=0.3 and 0.7.

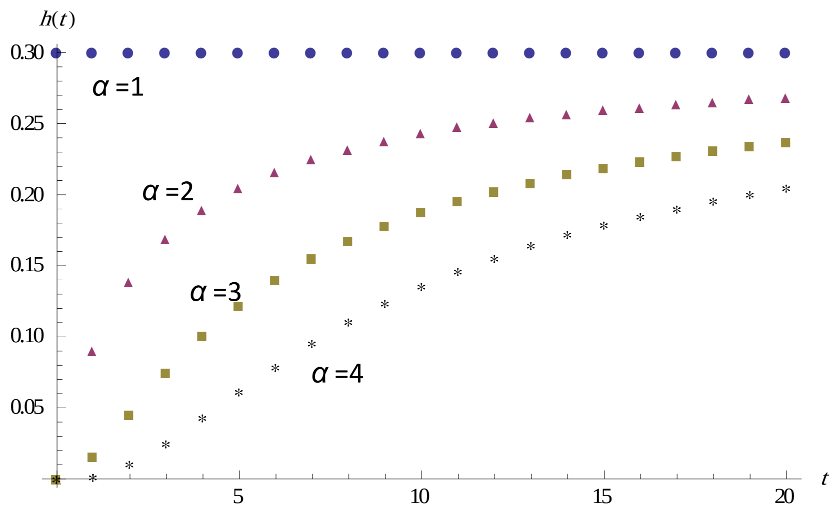


Figure 4. FR of DGD(α ,0.7)

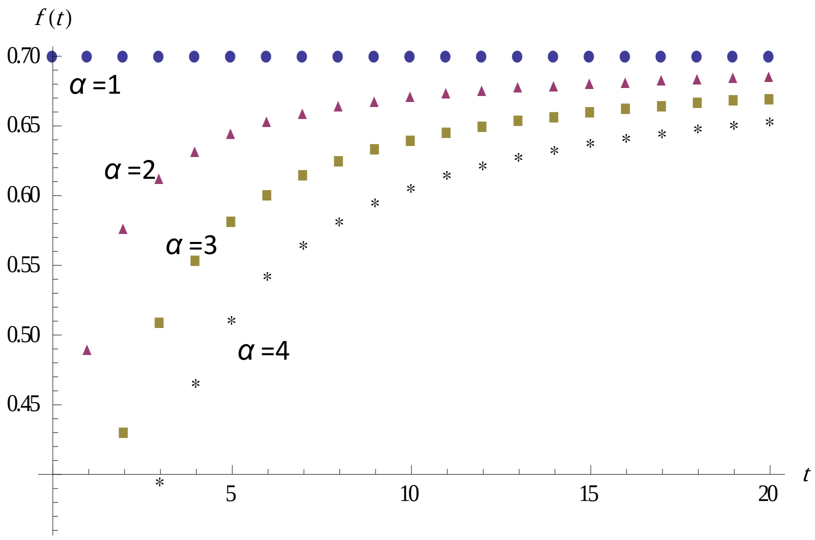


Figure 5. FR of DGD(α ,0.3)

Theorem 2: The $DGD(\alpha, q)$ is a unimodal distribution and the mode is given by $x^* = \left[\frac{1-\alpha}{\ln q}\right]$, here $[x]$ is the nearest integer to x.

Proof: This result follows immediately from the log-concavity of the DGD. The mode is obtained by solving $\frac{\partial}{\partial x} p_x = 0$ or

$C\,p^{\alpha}\,x^{\alpha-2}\,q^{x}[(\alpha-1)+x\ln q]=0$, which leads to mode of DGD is $x^{*}=\dfrac{1-\alpha}{\ln q}$.

Characterizations of unimodal discrete distribution are discussed in Abouammoh and Mashhour (1981). It is noted that the pmf of the $DGD(\alpha,q)$ can be characterized by similar relations.

Next, we are going to define a dual property of the failure rate, known by, the reversed failure rate.

Definition 2: The probability that an item survived at time x-1 given that it has failed before x+1 is defined to be the reversed failure rate. Formally,

$$r(x)=p(X>x-1/X\le x)=\frac{p(X=x)}{p(X\le x)}=\frac{p_x}{F(x)} \quad . \tag{3.9}$$

The reversed failure rate of the $DGD(\alpha,q)$ for $\alpha=2,3$ are given by

$$r(x)=\frac{x(1-q)^{2}}{q\,(qx+q^{-x}-x-1)} \tag{3.10}$$

$$r(x)=\frac{x^{2}\,(1-q)^{3}}{q[q^{1-x}+q^{-x}-q^{2}x^{2}-(1+x)^{2}-q(1-2x-2x^{2})]}\ , \tag{3.11}$$

respectively. From figure (6) it is observed that for all the values of α the $DGD(\alpha,q)$ has a decreasing reversed failure rate.

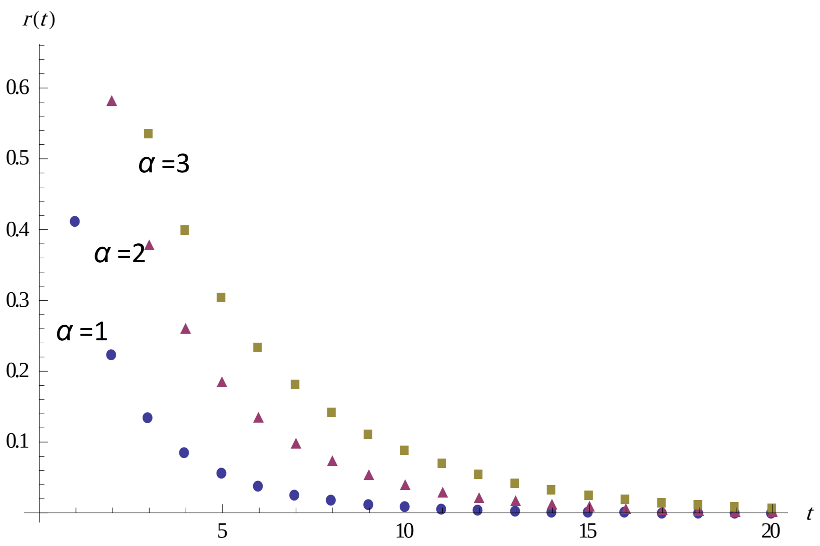


Figure 6. Reversed failure rate of DGD(α ,0.7)

3.4 The size-biased version of the DGD(α,q)

Assume that the lifetimes of a given sample of items follow the $DGD(\alpha, q)$ and the items are selected according to their life lengths means they are not equally likely to be selected then these lifetimes will have a size-biased $DGD(\alpha, q)$

Theorem 3: The k^{th} size-biased $DGD(\alpha, q)$ is a $DGD(\alpha + k, q)$, k=0,1,2,..

Proof: The k^{th} size-biased $DGD(\alpha, q)$ is given by:

$$\frac{x^k p_x}{E(X^k)} = \left(\frac{p^\alpha x^{\alpha+k-1} q^x}{q \sum_{i=0}^{\alpha-2} A(\alpha-1, i) q^i} \right) \left(\frac{p^k \sum_{i=0}^{\alpha-2} A(\alpha-1, i) q^i}{\sum_{i=0}^{\alpha+k-2} A(\alpha+k-1, i) q^i} \right)$$

$$= \frac{1}{q \sum_{i=0}^{\alpha+k-2} A(\alpha+k-1, i) q^i} p^{\alpha+k} x^{\alpha+k-1} q^x$$

Which is the pmf of the $DGD(\alpha + k, q)$.

Remark 4: When k=0 the weighted distribution of the $DGD(\alpha,q)$ reduces to the ordinary $DGD(\alpha,q)$, and when k=1 the weighted distribution becomes the length-biased distribution. Since the length-biased version of the $DGD(\alpha,q)$ ($CGD(\alpha,B)$) is again $DGD(\alpha+1,q)$ ($CGD(\alpha+1,\beta)$) with only a slight parameter shift, the detection of length-biased sampling using any test will be difficult in this case.

3.5 Relationship with other distributions

From the above discussion it is clear that the $DGD(\alpha,q)$ reduced to the G(q) when α=1. Furthermore, the next theorem will show that the $DGD(\alpha,q)$ appears to be a special linear combination of negative binomial distributions where the coefficients sum to one.

Theorem 4:

For $\alpha \geq 2$, $$DGD(\alpha,q) = \frac{1}{q\sum_{i=0}^{\alpha-2} A(\alpha-1,i)\, q^i} \sum_{i=1}^{\alpha} [a_i\, NB(i,q)] \qquad (5.1)$$

Where, $\sum_{i=1}^{\alpha} a_i = q\sum_{i=0}^{\alpha-2} A(\alpha-1,i) q^i$

Proof: The proof is carried out using the mathematical induction

1- $DGD(2,q) = \frac{1}{q}[NB(2) - pG(q)]$, where $1-p=q$.

2- The induction hypothesis, assume that (5.1) is true for some α=k

Therefore, $$DGD(k,q) = \frac{1}{q\sum_{i=0}^{k-2} A(k-1,i)\, q^i} \sum_{i=1}^{k} [a_i\, NB(i,q)],$$

where $\sum_{i=1}^{k} a_i = q\sum_{i=0}^{k-2} A(k-1,i) q^i$.

3- Now we show (5.1) is true for α=k+1, we need to prove that

$$DGD(k+1,q)=\frac{1}{q\sum_{i=0}^{k-1}A(k,i)q^i}\sum_{i=1}^{k+1}\left[b_i\ NB(i,q)\right],$$

where $\sum_{i=1}^{k+1}b_i=q\sum_{i=0}^{k-1}A(k,i)q^i$.

Now, using theorem 3 along with the induction hypothesis we get

$$DGD(k+1,q)=\frac{x\,DGD(k,q)}{E(X)}$$

$$=x\ \frac{p\sum_{i=0}^{k-2}A(k-1,i)q^i}{\sum_{i=0}^{k-1}A(k,i)q^i}\ \frac{\sum_{i=1}^{k}a_i\ NB(i,q)}{q\sum_{i=0}^{k-2}A(k-1,i)q^i}$$

$$=\frac{1}{q\sum_{i=0}^{k-1}A(k,i)q^i}\ x\,p\sum_{i=1}^{k}a_i\ NB(i,q)$$

But since, $\sum_{i=1}^{k}a_i=q\sum_{i=0}^{k-2}A(k-1,i)q^i$,

the following holds, $x\,p\sum_{i=1}^{k}a_i\ NB(i,q)=\sum_{i=1}^{k+1}b_i\ NB(i,q)$,

where $\sum_{i=1}^{k+1}b_i=q\sum_{i=0}^{k-1}A(k,i)q^i$.

Remark 5: As special cases of the above theorem we give the following linear combinations:

a) $DGD(1,q)=NB(1,q)=G(q)$.

b) $DGD(2,q)=\frac{1}{q}[NB(2,q)-pG(q)]$.

c) $DGD(3,q)=\frac{1}{q(1+q)}[2NB(3,q)-3pNB(2,q)+p^2G(q)]$.

d) $DGD(4,q) = \dfrac{1}{q(1+4q+q^2)}[6NB(4,q) - 12pNB(3,q) + 7p^2NB(2,q) - p^3G(q)].$

In fact the $DGD(\alpha,q)$ can cover in shape the geometric, negative binomial and various types of Weibull distributions for different values of the shape parameter α, see for example figure (7).

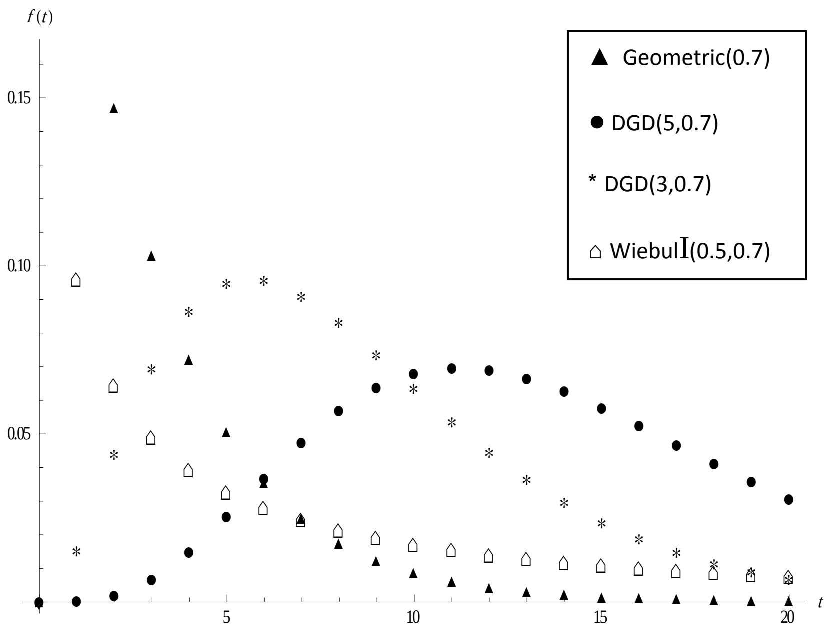


Figure 7. pmf of the geometric, DGD and Weibull type I when q= 0.7

Next, we introduce some finite mixtures of the $DGD(\alpha,q)$. A well known finite mixture of the continuous gamma is the Lindley distribution (1958). Here we present a generalization of the discrete Lindley.

3.6 Generalized discrete Lindley distribution

3.6.1 Basic definition

The pdf of the usual continuous Lindley distribution is given by

$$f(x)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} \;\; ; \;\; \theta>0 \;\; , x>0 \tag{6.1}$$

This distribution was introduced by Lindley (1958) by mixing the exponential and the continuous gamma distributions. Formally, the finite mixture is given by

$$f(x)=p_1\exp(\theta)+p_2 CGD(2,\theta) \;\; ,$$

where $p_1=\dfrac{\theta}{1+\theta}$ and $p_2=\dfrac{1}{1+\theta}$

Using similar manner, the generalized discrete Lindley distribution is defined by mixing G(q) and the $DGD(\alpha,q)$.

Definition 3: A discrete non-negative random variable X, is said to have a generalized discrete Lindley distribution with parameters (α,q), denoted by $GDLD(\alpha,q)$, if its pmf is given by

$$p(\alpha,q)=\frac{p^{\alpha}q^{x}}{1+p^{\alpha-1}}\left[1+\frac{x^{\alpha-1}}{q\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^{i}}\right] \;\; ; \;\; x=0,1,2..... \tag{6.2}$$

Here $\alpha=2,3,4...$ and $0<q<1$

The form of the finite mixture that led to the $GDLD(\alpha,q)$ is given by

$$GDLD(\alpha,q)=p_1G(q)+p_2DGD(\alpha,q) \;\; ,$$

where $p_1=\dfrac{p^{\alpha-1}}{1+p^{\alpha-1}}$ and $p_2=\dfrac{1}{1+p^{\alpha-1}}$

Note that, p_1 and p_2 are functions of p and α which are the parameters of the $DGD(\alpha,q)$, therefore, the invariance property of the mle's of p and α leads to the mle's of the weights p_1 and p_2.

The following are special cases of the $GDLD(\alpha, q)$ when α=2,3,4 :

$$p(2,q) = \frac{p^2 q^x}{1+p}\left[1+\frac{x}{q}\right] \tag{6.3}$$

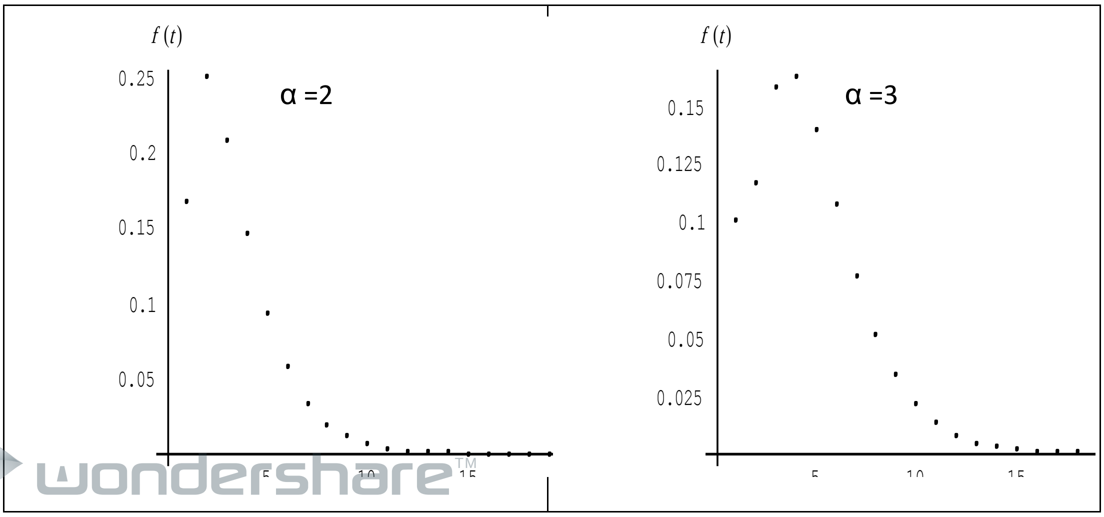
$$p(3,q) = \frac{p^3 q^x}{1+p^2}\left[1+\frac{x^2}{q(1+q)}\right] \tag{6.4}$$

$$p(4,q) = \frac{p^4 q^x}{1+p^3}\left[1+\frac{x^3}{q(1+4q+q^2)}\right] \quad , \tag{6.5}$$

respectively.

Remark 6: The pmf given in (6.3) can be considered as the discrete counter-part of the continuous Lindley given in (6.1).

Usually plots of the pmf for different values of the underlying parameters give more explanations for the shape of the distribution. Figures (8), (9) and (10) show the pmf of the $GDLD(\alpha, q)$ for different values of the parameters.



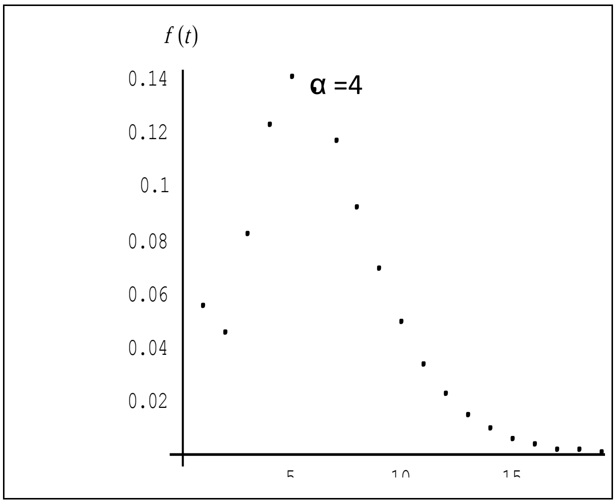


Figure 8. pmf of GDLD(α ,0.5)

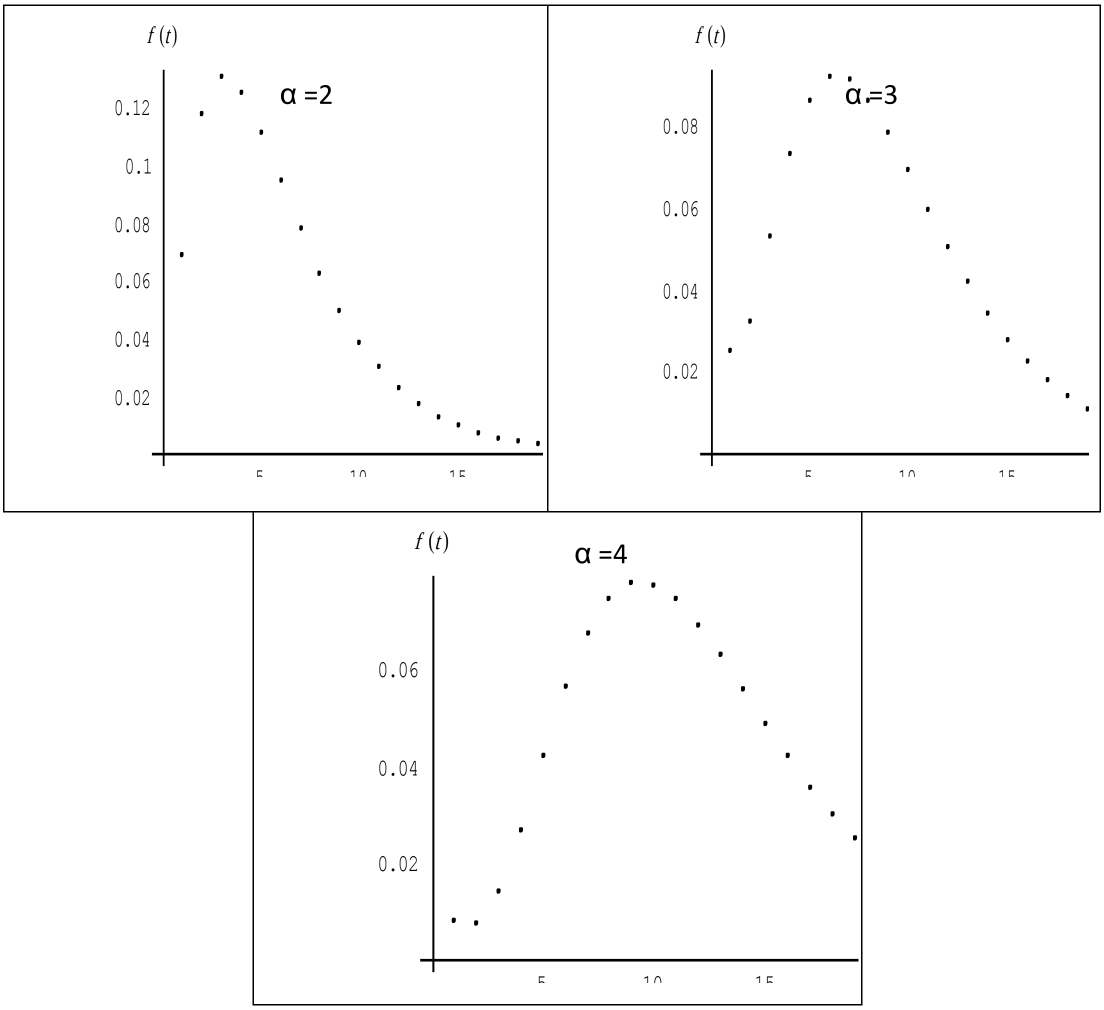


Figure 9. pmf of GDLD(α ,0.7)

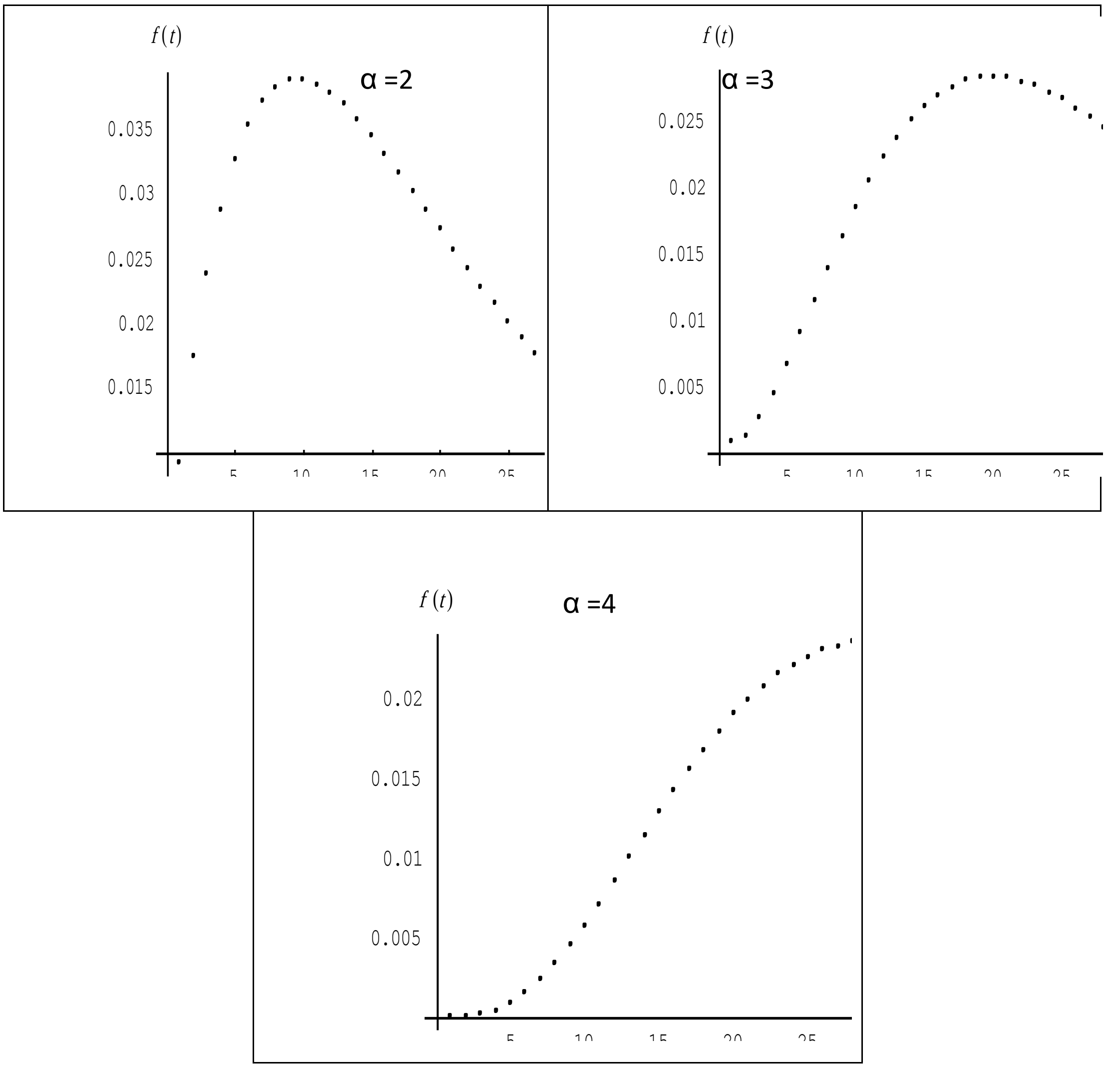


Figure 10. pmf of GDLD(α ,0.9)

It is clear that the $GDLD(\alpha,q)$ becomes approximately symmetric as α increases. This means that the skewness decreases as α increases. Also one can note that the peak of the $GDLD(\alpha,q)$ becomes less sharper while the tail becomes shorter as α increases. This means that the kurtosis decreases as α increases.

3.6.2 Statistical properties of the GDLD(α,q)

In this section the main statistical properties of the $GDLD(\alpha,q)$ are obtained.

(i) Using the k^{th} moment of both the G(q) and the $DGD(\alpha,q)$ lead to the k^{th} moment of the $GDLD(\alpha,q)$ which is given by

$$E(X^k)=\frac{p^{\alpha-1}}{1+p^{\alpha-1}}\left[\frac{q\sum_{i=0}^{k-1}A(k,i)q^i}{p^k}\right]+\frac{1}{1+p^{\alpha-1}}\left[\frac{\sum_{i=0}^{\alpha+k-2}A(\alpha+k-1,i)q^i}{p^k\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i}\right] \tag{6.6}$$

(ii)The mean and the second moment can be given by

$$E(X)=\frac{p^{\alpha-1}}{1+p^{\alpha-1}}\left[\frac{q}{p}\right]+\frac{1}{1+p^{\alpha-1}}\left[\frac{\sum_{i=0}^{\alpha-1}A(\alpha,i)q^i}{p\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i}\right] \tag{6.7}$$

$$E(X^2)=\frac{p^{\alpha-1}}{1+p^{\alpha-1}}\left[\frac{q(1+q)}{p^2}\right]+\frac{1}{1+p^{\alpha-1}}\left[\frac{\sum_{i=0}^{\alpha}A(\alpha+1,i)q^i}{p^2\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i}\right], \tag{6.8}$$

respectively.

(iii) The variance is evaluated based on (6.8) and (6.7) by

$$V(X)=E(X^2)-[E(X)]^2 \tag{6.9}$$

Next, we present the mean, the second moment and the variance of the $GDLD(2,q)$

$$E(X)=\frac{q(1+p)+1}{p(1+p)} \tag{6.10}$$

$$E(X^2)=\frac{pq(1+q)+1+4q+q^2}{p^2(1+p)} \tag{6.11}$$

$$V(X)=\frac{p+2q+3pq+p^2q}{p^2(1+p)^2}, \tag{6.12}$$

respectively.

Whereas, for the $GDLD(3,q)$ the mean and the second moment are given by

$$E(X)=\frac{p^2q(1+q)+1+4q+q^2}{(p^2+1)(1+q)p} \tag{6.13}$$

$$E(X^2)=\frac{p^2q(1+q)^2+1+11q+11q^2+q^3}{(p^2+1)(1+q)p^2}, \tag{6.14}$$

respectively.

Also, the moment generating function (mgf) of the $GDLD(\alpha,q)$ is found to be:

$$E(e^{tX})=\frac{p^{\alpha-1}}{1+p^{\alpha-1}}\left[\frac{1-q}{1-e^tq}\right]+\frac{1}{1+p^{\alpha-1}}\left[\frac{e^t\,p^{\alpha}[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)(qe^t)^i]}{(1-qe^t)^{\alpha}[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i]}\right] \tag{6.15}$$

Hence, the factorial mgf of the $GDLD(\alpha,q)$ is given by:

$$E(s^X)=\frac{p^{\alpha-1}}{1+p^{\alpha-1}}\left[\frac{1-q}{1-sq}\right]+\frac{1}{1+p^{\alpha-1}}\left[\frac{s\,p^{\alpha}[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)(qs)^i]}{(1-qs)^{\alpha}[\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i]}\right] \tag{6.16}$$

Remark 7: Unlike the Poisson and the NB(r,q), the GDLD(α,q) can handle both over and under-dispersed data.

Using (6.10) and (6.12)

$V(X)-E(X)=\dfrac{-1+4q+2q^2-4q^3+q^4}{(2-3q+q^2)^2}=(>)(<)0$ if and only if $q=(>)(<)0.2346$.

This means that the $GDLD(2,q)$ is equi-dispersed, over-dispersed and under-dispersed for $q=(>)(<)0.2346$, respectively.

Similarly, for the $GDLD(3,q)$

$$V(X)-E(X)=\frac{q^4+8q^3+4q^2-1}{(1-q^2)^2}=(>)(<)0 \text{ if and only if } q=(>)(<)0.3744.$$

This means that the $DGD(3,q)$ is equi-dispersed, over-dispersed and under-dispersed for $q=(>)(<)0.3744$, respectively.

3.6.3 Reliability Properties of the GDLD(α,q)

In this section the basic reliability properties of the $GDLD(2,q)$ are investigated.

The cdf, sf and FR of the $GDLD(2,q)$ are given by:

$$F(x)=\frac{2-q+q^{2+x}+q^{1+x}(x-1)-q^x(1+x)}{2-q} \tag{6.17}$$

$$S(x)=\frac{q^x(1-q^2-q(x-1)+x)}{2-q} \tag{6.18}$$

$$h(x)=\frac{-p^2(q+x)}{q^2+q(x-2)-x}\quad , \tag{6.19}$$

respectively.

Whereas, for the $GDLD(3,q)$, the cdf, sf and FR are given by:

$$F(x)=\frac{2-q^2+q^3+q^{x+3}-q^{x+4}-q^x(1+x)^2-q^{x+2}(x^2-1)+2q^{x+1}(x^2+x-1)}{2-q^2+q^3} \tag{6.20}$$

$$S(x)=\frac{q^x(-q^3+q^4+(1+x)^2+q^2(x^2-1)-2q(x^2+x-1))}{2-q^2+q^3} \tag{6.21}$$

$$h(x)=\frac{(1-q)^3(q+q^2+x^2)}{-q^3+q^4+q^2(x-2)x+x^2+q(2+2x-2x^2)}, \tag{6.22}$$

respectively.

Figures (11), (12) and (13) represent the graphs of the FR for different values of α when q=0.3, 0.5 and 0.7.

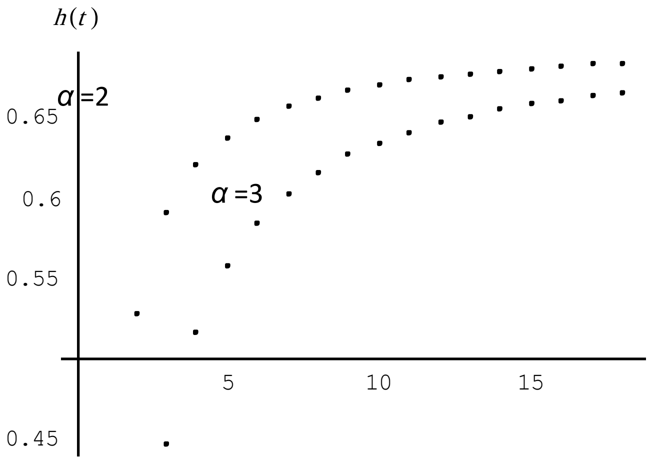


Figure 11. FR of GDLD(α ,0.3)

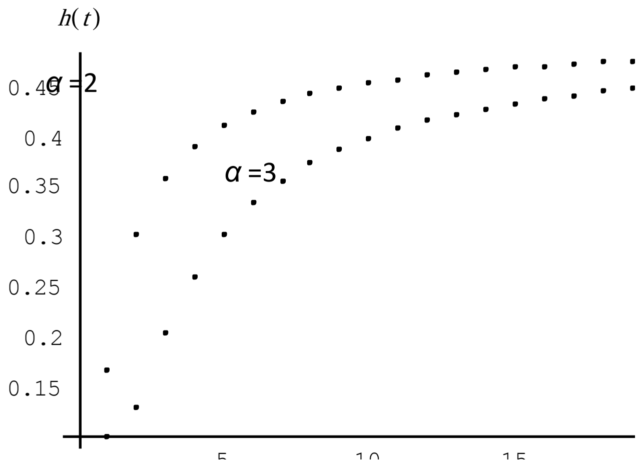


Figure 12. FR of GDLD(α ,0.5)

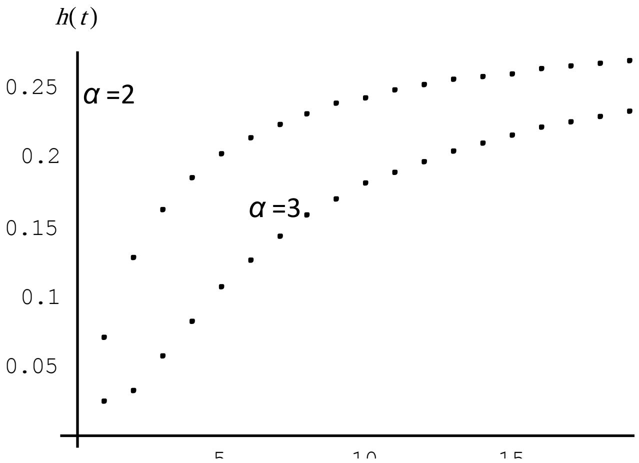


Figure 13. FR of GDLD(α ,0.7)

3.7 Generalized discrete Lindley distribution II

In this section a new discrete generalization of the Lindley distribution is introduced and is given the name type II.

3.7.1 Basic definition

Definition 4: A discrete non-negative random variable X, is said to have a generalized discrete Lindley distribution II with parameters (α,q), denoted by $GDLD_2(\alpha,q)$, if its pmf is given by

$$p(\alpha,q)=\frac{p^{\alpha+1}q^{x}}{1+p}\left[\frac{x^{\alpha-1}}{q\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^{i}}+\frac{x^{\alpha}}{q\sum_{i=0}^{\alpha-1}A(\alpha,i)q^{i}}\right]\quad ;\quad x=0,1,2..... \tag{7.1}$$

Here $\alpha=2,3,4...$ and $0<q<1$

The form of the finite mixture that led to the $GDLD_2(\alpha,q)$ is given by

$$GDLD_2(\alpha,q)=p_1DGD(\alpha,q)+p_2DGD(\alpha+1,q)\ ,$$

where $p_1=\dfrac{p}{1+p}$ and $p_2=\dfrac{1}{1+p}$

Note that, p_1 and p_2 are functions of p which is the parameter of the $DGD(\alpha,q)$, therefore, the invariance property of the mle of p leads to the mle's of the weights p_1 and p_2.

The following are special cases of the $GDLD_2(\alpha,q)$ when α=2,3 :

$$p(2,q)=\frac{p^3q^x}{1+p}\left[\frac{x}{q}+\frac{x^2}{q(1+q)}\right] \tag{7.2}$$

$$p(3,q)=\frac{p^4q^x}{1+p^2}\left[\frac{x^2}{q(1+q)}+\frac{x^3}{q(1+4q+q^2)}\right] \tag{7.3}$$

respectively. Figures (14), (15) and (16) show the pmf of the $GDLD_2(\alpha,q)$ for different values of the parameters.

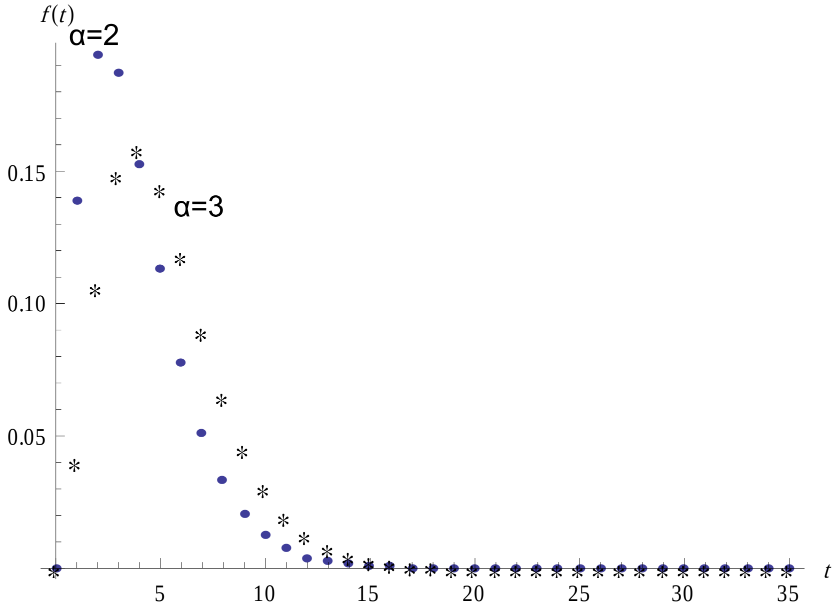


Figure 14. pmf of $GDLD_2(\alpha,0.5)$

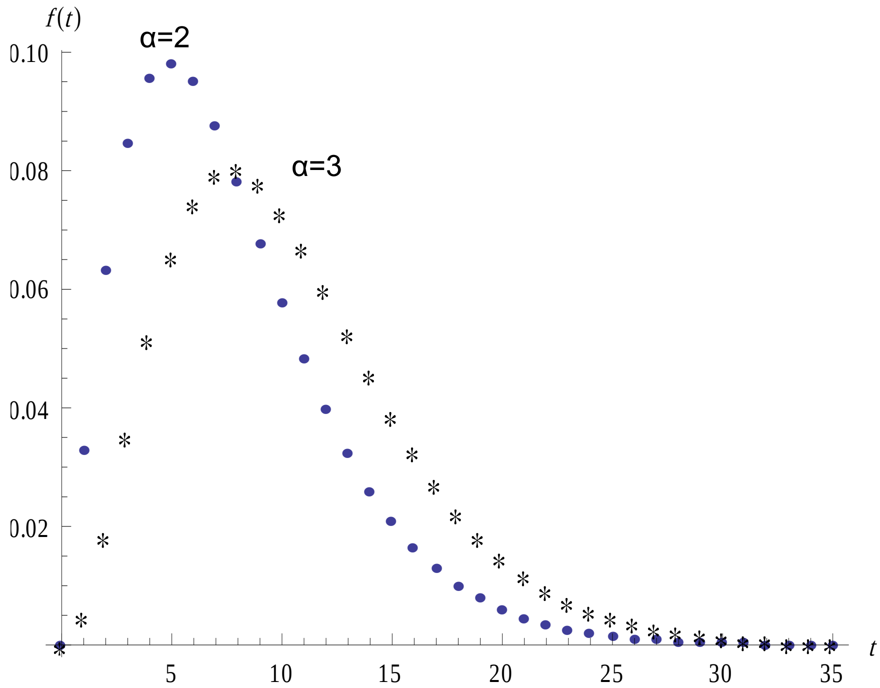


Figure 15. pmf of $GDLD_2(\alpha, 0.7)$

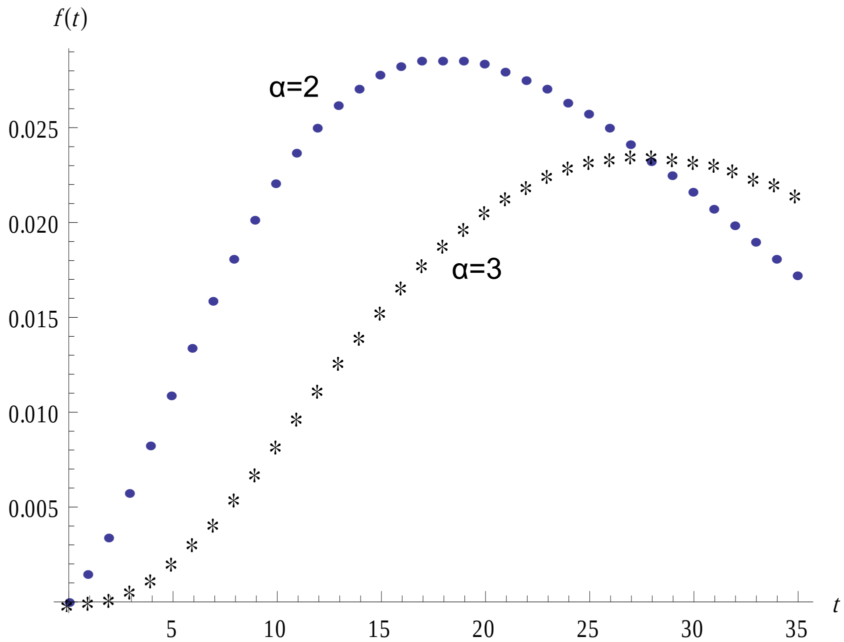


Figure 16. pmf of $GDLD_2(\alpha, 0.9)$

From the above figures one can see that the $GDLD_2(\alpha, q)$ is right skewed and the kurtosis decreases as α or q increases.

3.7.2 Statistical properties of the $GDLD_2(\alpha, q)$

In this section the main statistical properties of the $GDLD_2(\alpha, q)$ are obtained.

(i) Using the k^{th} moment of both the $DGD(\alpha, q)$ and the $DGD(\alpha+1, q)$ lead to the k^{th} moment of the $GDLD_2(\alpha, q)$ which is given by

$$E(X^k) = \frac{p}{1+p}\left[\frac{\sum_{i=0}^{\alpha+k-2} A(\alpha+k-1, i)q^i}{p^k \sum_{i=0}^{\alpha-2} A(\alpha-1, i)q^i}\right] + \frac{1}{1+p}\left[\frac{\sum_{i=0}^{\alpha+k-1} A(\alpha+k, i)q^i}{p^k \sum_{i=0}^{\alpha-1} A(\alpha, i)q^i}\right] \quad (7.4)$$

(ii)The mean and the second moment can be given by

$$E(X) = \frac{p}{1+p}\left[\frac{\sum_{i=0}^{\alpha-1} A(\alpha, i)q^i}{p \sum_{i=0}^{\alpha-2} A(\alpha-1, i)q^i}\right] + \frac{1}{1+p}\left[\frac{\sum_{i=0}^{\alpha} A(\alpha+1, i)q^i}{p \sum_{i=0}^{\alpha-1} A(\alpha, i)q^i}\right] \quad (7.5)$$

$$E(X^2) = \frac{p}{1+p}\left[\frac{\sum_{i=0}^{\alpha} A(\alpha+1, i)q^i}{p^2 \sum_{i=0}^{\alpha-2} A(\alpha-1, i)q^i}\right] + \frac{1}{1+p}\left[\frac{\sum_{i=0}^{\alpha+1} A(\alpha+2, i)q^i}{p^2 \sum_{i=0}^{\alpha-1} A(\alpha, i)q^i}\right], \quad (7.6)$$

respectively.

(iii) The variance is evaluated based on (7.5) and (7.6) by

$$V(X) = E(X^2) - [E(X)]^2$$

Next, we present the mean, the second moment and the variance of the $GDLD_2(2, q)$

$$E(X) = \frac{2+5q-q^3}{2-q-2q^2+q^3} \quad (7.7)$$

$$E(X^2) = \frac{-2-13q+2q^2+q^3}{(q-2)(1-q)^2} \quad (7.8)$$

$$V(X)=\frac{2q(6+5q-3q^2-3q^3+q^4)}{(1+q)^2(2-3q+q^2)^2}, \tag{7.9}$$

respectively.

Remark 7: The $GDLD_2(2,q)$ can handle both over and under-dispersed data.

Using (7.7) and (7.9)

$V(X)-E(X)=\frac{-4+4q+19q^2+4q^3-12q^4+q^6}{(1+q)^2(2-3q+q^2)^2}=(>)(<)0$ if and only if $q=(>)(<)0.3664$.

This means that the $GDLD_2(2,q)$ is equi-dispersed, over-dispersed and under-dispersed for $q=(>)(<)0.3664$, respectively.

3.7.3 Reliability Properties of the $GDLD_2(\alpha,q)$

In this section the basic reliability properties of the $GDLD_2(2,q)$ are investigated.

The cdf and the FR of the $GDLD_2(2,q)$ is given by:

$$F(x)=\frac{-2-q+q^2+xq^{x+3}+(1-3x-2x^2)q^{x+1}-(1+x-x^2)q^{x+2}+(2+3x+x^2)q^x}{q^2-q-2} \tag{7.10}$$

$$h(x)=\frac{(1-q)^3x(1+q+x)}{(x-1)q^3+x(x+1)+(2+x-2x^2)q+(1-3x+x^2)q^2}, \tag{7.11}$$

respectively.

Figure (17) represent the graph of the FR of the $GDLD_2(2,q)$ for different values of q. It is clear that the $GDLD_2(2,q)$ has an IFR.

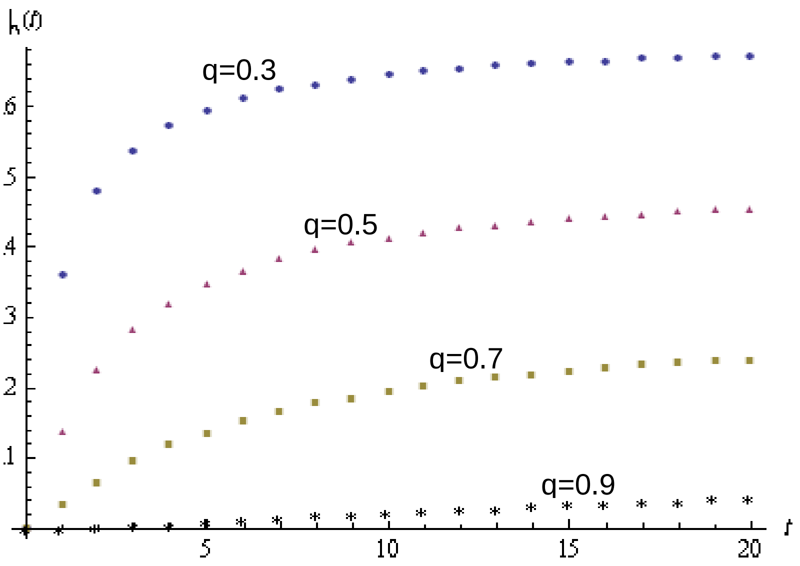


Figure 17. FR of $GDLD_2(2,q)$

3.8 Generalized discrete gamma distribution

Here, the $DGD(\alpha,q)$ is generalized by considering a third parameter.

3.8.1 Basic definition

The pdf of the generalized continuous gamma is given by

$$f(x)=\frac{\theta\beta^{\alpha}}{\Gamma\left(\frac{\alpha}{\theta}\right)}x^{\alpha-1}e^{-(\beta x)^{\theta}}\ ;\ x\geq 0\ ,\ \alpha>0\ ,\ \beta>0\ ,\theta>0\ .$$

If θ= 1, the generalized gamma reduces to the gamma distribution. And If α=θ then the generalized gamma distribution becomes the Wiebull distribution. For more details about the generalized gamma distribution see Stacy (1962).

Definition 5: A discrete non-negative random variable X, is said to have a generalized discrete gamma distribution with parameters (α,β,q), denoted by $GDGD(\alpha,\beta,q)$, if its pmf is given by

$$p_x = p(X = x) = C\, x^{\alpha-1} q^{\beta x}; \quad x = 0,1,2....., \quad \alpha, \beta = 1,2......, \quad 0 < q < 1 \qquad (8.1)$$

where $c = (1 - q^{\beta})$ *if* $\alpha = 1$ and $c = \dfrac{(1-q^{\beta})^{\alpha}}{q^{\beta}\,[\sum_{i=0}^{\alpha-2} A(\alpha-1, i)\, q^{i\beta}]}$ *for* $\alpha \geq 2$.

Remark 8: Since $GDGD(\alpha, \beta, q) = DGD(\alpha, q^{\beta})$, all the statistical and reliability properties listed above of the $DGD(\alpha, q)$ with replacing q by q^{β} are also the properties of the $GDGD(\alpha, \beta, q)$.

Theorem 5: If $X \sim GDGD(\alpha, \beta, q)$, then X has an IFR for $x \geq 1, \alpha \geq 2, \beta \geq 1$.

Proof: Consider the function

$$b(x, \alpha) = \frac{p(X = x+1)}{p(X = x)} = q^{\beta}\left(\frac{x+1}{x}\right)^{\alpha-1}$$

Its derivative is given by

$$\frac{db(x,\alpha)}{dx} = -q^{\beta}\,\frac{\alpha-1}{x^2}\left(\frac{x+1}{x}\right)^{\alpha-2} < 0 \qquad \text{for } \alpha \geq 2, \beta \geq 1 \,.$$

Note that $b(x, \alpha)$ is a decreasing function in x for $x \geq 1, \alpha \geq 2, \beta \geq 1$ thus the $GDGD(\alpha, \beta, q)$ is log-concave and hence the FR is increasing.

Theorem 6: The $GDGD(\alpha, \beta, q)$ is a unimodal distribution and the mode is given by $x^* = \left[\dfrac{1-\alpha}{\ln q^{\beta}}\right]$

Proof: This result follows immediately from the log-concavity of the $GDGD(\alpha, \beta, q)$. The mode is obtained by solving $\dfrac{\partial}{\partial x} p_x = 0$ or $C\, x^{\alpha-2} q^{\beta x}[(\alpha - 1) + x \ln q^{\beta}] = 0$, which leads to mode of $GDGD(\alpha, \beta, q)$ that is $x^* = \dfrac{1-\alpha}{\ln q^{\beta}}$.

3.8.2 The size-biased version of the GDGD(α,β,q)

assume that the lifetimes of a given sample of items follow a $GDGD(\alpha,\beta,q)$ and the items are selected according to their life lengths means they are not equally likely to be selected then these lifetimes will have a size-biased $GDGD(\alpha,\beta,q)$

Theorem 7: The k^{th} size-biased $GDGD(\alpha,\beta,q)$ is a $GDGD(\alpha+k,\beta,q)$, k=0,1,2,..

Proof: The k^{th} size-biased $GDGD(\alpha,\beta,q)$ is given by:

$$\frac{x^k\,p_x}{E(X^k)} = \left(\frac{(1-q^\beta)^\alpha x^{\alpha+k-1} q^{\beta x}}{q^\beta \sum_{i=0}^{\alpha-2} A(\alpha-1,i) q^{i\beta}}\right)\left(\frac{(1-q^\beta)^k \sum_{i=0}^{\alpha-2} A(\alpha-1,i) q^{i\beta}}{\sum_{i=0}^{\alpha+k-2} A(\alpha+k-1,i) q^{i\beta}}\right)$$

$$= \frac{1}{q^\beta \sum_{i=0}^{\alpha+k-2} A(\alpha+k-1,i) q^{i\beta}} (1-q^\beta)^{\alpha+k} x^{\alpha+k-1} q^{\beta x}$$

Which is the pmf of the $GDGD(\alpha+k,\beta,q)$.

Chapter 4

Parameter Estimation

In this chapter two methods of estimation namely, the maximum likelihood method and method of moments are used to estimate the parameters of the $DGD(\alpha,q)$, $GDLD(\alpha,q)$ and $GDLD_2(\alpha,q)$. Applications in real-life data will be given and their goodness-of-fit to the distributions using the Kolmogorov-Smirnov test are examined. Moreover, graphical analysis will be used in conjunction with the formal test to confirm the results. In this piece of work we restrict our effort to point estimation and investigate some properties of estimators by simulation. In fact estimation of the underlying parameters of any probability model is necessary for further real-life applications. Also, samples generation is discussed briefly for each of the distributions.

4.1 Estimation of the DGD parameters

4.1.1 Generating sample from the $DGD(\alpha,q)$

Samples of the $DGD(\alpha,q)$ can be generated by solving the following equations for t_i

$$F(t_i) - u_i = 0, \tag{1.1}$$

where $\{u_i : i = 1,2,....n\}$ is a sample generated from the uniform distribution. Since t_i might be negative and non-integer we need to consider $x_i = [y_i]$ *where* $y_i = |t_i|$. For example, samples of the $DGD(1,q)$, $DGD(2,q)$ and $DGD(3,q)$ are generated by solving the following equations for t_i

$$1-q^{t_i+1}-u_i=0 \tag{1.2}$$

$$1-q^{t_i}-t_i\,q^{t_i}+t_i\,q^{t_i+1}-u_i=0 \tag{1.3}$$

$$\frac{1+q-(1+t_i)^2 q^{t_i}-(1-2t_i-2t_i^{\,2})q^{t_i+1}-t_i^{\,2}\,q^{t_i+2}}{(1+q)}-u_i=0 \quad , \tag{1.4}$$

respectively.

4.1.2 Maximum likelihood method

In this subsection the maximum likelihood estimation of the $DGD(\alpha,q)$ parameters when both α and q are unknown is given by the complicated likelihood function which is given by:

$$L(\alpha,q)=\frac{(1-q)^{n\alpha}}{\left(\sum_{i=0}^{\alpha-2}A(\alpha-1,i)q^i\right)^n}\left(\prod_{i=1}^{n}x_i\right)^{\alpha-1} q^{\sum_{i=1}^{n}x_i-n} \quad . \tag{1.5}$$

Therefore, this method is to be considered to estimate the parameter q when α is known.

The Likelihood function of the DGD(1,q) which is the G(q) for a sample of size n is given by

$$L(q)=(1-q)^n\,q^{\sum_{i=1}^{n}x_i} \quad .$$

Thus, the log likelihood function is

$$\ln L(q)=n\ln(1-q)+(\sum_{i=1}^{n}x_i)\ln(q)\,. \tag{1.6}$$

Differentiating (1.6) with respect to q and then equating to zero, gives the mle of q that is:

$$\hat{q} = \frac{\sum_{i=1}^{n} x_i}{n + \sum_{i=1}^{n} x_i} . \qquad (1.7)$$

The Likelihood function of the DGD(2,q) for a sample of size n is given by

$$L(q) = (1-q)^{2n} \left(\prod_{i=1}^{n} x_i \right) q^{\sum_{i=1}^{n} x_i - n} .$$

Thus, the log likelihood function is

$$\ln L(q) = 2n \ln(1-q) + \sum_{i=1}^{n} \ln(x_i) + \left(\sum_{i=1}^{n} x_i - n\right) \ln(q) . \qquad (1.8)$$

The mle of q is the value of q that maximizes $\ln L(q)$. Differentiating (1.8) with respect to q and then equating to zero, gives the mle of q that is:

$$\hat{q} = \frac{\sum_{i=1}^{n} x_i - n}{\sum_{i=1}^{n} x_i + n} . \qquad (1.9)$$

Here, $\frac{\partial^2 \ln l(q)}{\partial q^2} = \frac{-2n}{(1-q)^2} - \frac{n(\bar{x}-1)}{q^2}$.

Therefore, the $(1-\alpha)100\%$ CI of q is given by

$$\hat{q} \pm z_{\alpha/2} \sqrt{I^{-1}(q)} ,$$

where $I(q) = -E[\frac{\partial^2 \ln l(q)}{\partial q^2}] = \frac{2n}{(1-q)^2} + \frac{n(\mu-1)}{q^2}$

The likelihood function of the DGD(3,q) for a sample of size n is given by:

$$L(q)=\frac{1}{(1+q)^n}\left(\prod_{i=1}^{n}x_i^2\right)(1-q)^{3n}\ q^{\sum_{i=1}^{n}x_i-n}.$$

Thus, the log likelihood function is

$$\ln L(q)=-n\ln(1+q)+3n\ln(1-q)+2\sum_{i=1}^{n}\ln(x_i)+(\sum_{i=1}^{n}x_i-n)\ln(q). \qquad (1.10)$$

Differentiating (1.10) with respect to q and then equating to zero gives the following equation:

$$-\frac{n}{1+q}-\frac{3n}{1-q}+\frac{\sum_{i=1}^{n}x_i-n}{q}=0.$$

This leads to the mle of q that is

$$\hat{q}=\frac{-2n\pm\sqrt{3n^2-\left(\sum_{i=1}^{n}x_i\right)^2}}{\sum_{i=1}^{n}x_i+n}. \qquad (1.11)$$

Here, $\frac{\partial^2 \ln l(q)}{\partial q^2}=\frac{n}{(1+q)^2}-\frac{3n}{(1-q)^2}-\frac{n(\bar{x}-1)}{q^2}$.

Therefore, the $(1-\alpha)100\%$ CI of q is given by

$$\hat{q}\pm z_{\alpha/2}\sqrt{I^{-1}(q)},$$

where $I(q)=-E[\frac{\partial^2 \ln l(q)}{\partial q^2}]=-\frac{n}{(1+q)^2}+\frac{3n}{(1-q)^2}+\frac{n(\mu-1)}{q^2}$

The likelihood function of the DGD(4,q) for a sample of size n is given by:

$$L(q)=\frac{1}{(1+4q+q^2)^n}\left(\prod_{i=1}^{n}x_i^3\right)(1-q)^{4n}\ q^{\sum_{i=1}^{n}x_i-n}.$$

Thus, the log likelihood function is

$$\ln L(q)=-n\ln(1+4q+q^2)+4n\ln(1-q)+3\sum_{i=1}^{n}\ln(x_i)+(\sum_{i=1}^{n}x_i-n)\ln(q) \qquad (1.12)$$

Differentiating (1.12) with respect to q and then equating to zero gives the following equation which will be solved numerically for q:

$$-\frac{n(4+2q)}{1+4q+q^2}-\frac{4n}{1-q}+\frac{\sum_{i=1}^{n}x_i-n}{q}=0. \qquad (1.13)$$

Following the same procedure, the equation that gives the mle of the parameter q of the DGD(5,q) is given by:

$$-\frac{n(11+22q+3q^2)}{1+11q+11q^2+q^3}-\frac{5n}{1-q}+\frac{\sum_{i=1}^{n}x_i-n}{q}=0. \qquad (1.14)$$

Corollary 1: For a given random sample from the DGD(2,q), $\bar{x}-1$ is a MVUE as well as consistent of the function $h(q)=\frac{2q}{1-q}$.

Proof:

$$\begin{aligned}\frac{\partial \ln L(q)}{\partial q}&=-\frac{2n}{1-q}+\frac{\sum_{i=1}^{n}x_i}{q}-\frac{n}{q}\\&=\frac{n}{q}[(\bar{x}-1)-\frac{2q}{1-q}]\\&=a(q)[T-h(q)].\end{aligned}$$

Note that a(q) is independent of the observations so based on the result of Cramer-Rao theorem, $T=\bar{x}-1$ is a MVUE of h(q). Moreover, the variance of $\bar{x}-1$ is given by

$$V(\bar{x}-1)=V(\bar{x})=\frac{h'(q)}{a(q)}=\frac{2q}{n(1-q)^2}.$$

Now, since $V(T)\to 0\, as\, n\to\infty$ then T is a consistent estimator of h(q)

Corollary 2: For a given random sample from the DGD(3,q), $\bar{x}-1$ is a MVUE of the function $h(q)=\frac{2q(2+q)}{(1-q)(1+q)}$.

Proof:

$$\frac{\partial \ln L(q)}{\partial q}=-\frac{n}{1+q}-\frac{3n}{(1-q)}+\frac{\sum_{i=1}^{n}x_i}{q}-\frac{n}{q}$$

$$=\frac{n}{q}\left[(\bar{x}-1)-q\left[\frac{3}{1-q}+\frac{1}{1+q}\right]\right]$$

$$=\frac{n}{q}\left[(\bar{x}-1)-\frac{2q(2+q)}{(1-q)(1+q)}\right]$$

$$=a(q)[T-h(q)].$$

Based on the result of Cramer-Rao theorem, $T=\bar{x}-1$ is a MVUE of h(q).

Corollary 3: For a given random sample from the DGD(4,q), $\bar{x}-1$ is a MVUE of the function $h(q)=\frac{2q(4+7q+q^2)}{(1-q)(1+4q+q^2)}$.

4.1.3 Simulation study

Here 1000 different samples of sizes 10, 20, 30, 50 and 100 are generated from the $DGD(\alpha,q)$ when α is known by using (1.1). The population value of q is considered to be 0.4, 0.5, 0.7 and 0.9. The maximum likelihood estimates of q which are the average estimates taken over the 1000 replicates are given in table 1, 2, 3, 4 and 5 also the mean square error and the bias of the estimates are calculated.

Table 1. mle of q for the DGD(1,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.1815	0.1978	0.2025	0.2074	0.2103
mse	0.0655	0.0505	0.0460	0.0413	0.0382
bias	-0.2185	-0.2022	-0.1975	-0.1926	-0.1897
(q=0.5)					
$\hat{q}_{mle}$	0.2929	0.3159	0.3231	0.3279	0.3320
mse	0.0665	0.04587	0.0395	0.0348	0.0308
bias	-0.2071	-0.1841	-0.1769	-0.1721	-0.1680
(q=0.7)					
$\hat{q}_{mle}$	0.5804	0.6011	0.6085	0.6147	0.6189
mse	0.0317	0.0179	0.0135	0.0102	0.0081
bias	-0.1196	-0.0989	-0.0915	-0.0853	-0.0811
(q=0.9)					
$\hat{q}_{mle}$	0.8760	0.8837	0.8862	0.8885	0.8897
mse	0.0029	0.0011	0.0007	0.0004	0.0002
bias	-0.0241	-0.0163	-0.0138	-0.0115	-0.0103

Table 2. mle of q for the DGD(2,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.1029	0.1250	0.1317	0.1396	0.1432
mse	0.1229	0.0918	0.0827	0.0740	0.0693
bias	-0.2971	-0.2750	-0.2683	-0.2604	-0.2568

(q=0.5)					
$\hat{q}_{mle}$	0.3009	0.3182	0.3234	0.3301	0.3332
mse	0.0638	0.0439	0.0384	0.0330	0.0300
bias	-0.1991	-0.1818	-0.1766	-0.1699	-0.1668
(q=0.7)					
$\hat{q}_{mle}$	0.6284	0.6381	0.6415	0.6454	0.6470
mse	0.0126	0.0072	0.0056	0.0042	0.0035
bias	-0.0716	-0.0619	-0.0585	-0.0546	-0.0530
(q=0.9)					
$\hat{q}_{mle}$	0.8890	0.8920	0.8930	0.8942	0.8947
mse	0.0008	0.0004	0.0002	0.0001	0.0001
bias	-0.0110	-0.0080	-0.0070	-0.0058	-0.0053

Table 3. mle of q for the DGD(3,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.2788	0.2605	0.2604	0.2641	0.2650
mse	0.0605	0.0267	0.0227	0.0204	0.0193
bias	-0.1212	-0.1395	-0.1396	-0.1359	-0.1350
(q=0.5)					
$\hat{q}_{mle}$	0.3930	0.3992	0.4015	0.4048	0.4060
mse	0.0193	0.0141	0.0123	0.0106	0.0097
bias	-0.1070	-0.1008	-0.0985	-0.0952	-0.0940
(q=0.7)					
$\hat{q}_{mle}$	0.6572	0.6622	0.6642	0.6665	0.6672
mse	0.0053	0.0031	0.0024	0.0018	0.0014
bias	-0.0428	-0.0378	-0.0358	-0.0335	-0.0328

(q=0.9)					
$\hat{q}_{mle}$	0.8930	0.8948	0.8955	0.8963	0.8966
mse	0.0004	0.0002	0.0001	0.0001	0.0000
bias	-0.0070	-0.0052	-0.0045	-0.0037	-0.0034

Table 4. mle of q for DGD(4,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.2991	0.3017	0.3033	0.3055	0.3059
mse	0.0156	0.0124	0.0112	0.0100	0.0095
bias	-0.1009	-0.0983	-0.0967	-0.0945	-0.0941
(q=0.5)					
$\hat{q}_{mle}$	0.4244	0.4283	0.4302	0.4327	0.4332
mse	0.0105	0.0076	0.0065	0.0055	0.0050
bias	-0.0756	-0.0717	-0.0698	-0.0673	-0.0668
(q=0.7)					
$\hat{q}_{mle}$	0.6671	0.6706	0.6718	0.6739	0.6743
mse	0.0034	0.0020	0.0016	0.0011	0.0009
bias	-0.0329	-0.0294	-0.0282	-0.0261	-0.0257
(q=0.9)					
$\hat{q}_{mle}$	0.8949	0.8962	0.8967	0.8973	0.8975
mse	0.0003	0.0001	0.0001	0.0001	0.0000
bias	-0.0051	-0.0038	-0.0033	-0.0027	-0.0025

Table 5. mle of q for DGD(5,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.3210	0.3232	0.3244	0.3264	0.3266
mse	0.0100	0.0079	0.0071	0.0062	0.0058
bias	-0.0790	-0.0768	-0.0756	-0.0736	-0.0734
(q=0.5)					
$\hat{q}_{mle}$	0.4414	0.4441	0.4455	0.4474	0.4479
mse	0.0069	0.0049	0.0042	0.0035	0.0031
bias	-0.0586	-0.0559	-0.0545	-0.0526	-0.0521
(q=0.7)					
$\hat{q}_{mle}$	0.6759	0.6784	0.6795	0.6810	0.6814
mse	0.0022	0.0013	0.0010	0.0007	0.0005
bias	-0.0241	-0.0216	-0.0205	-0.0190	-0.0186
(q=0.9)					
$\hat{q}_{mle}$	0.8960	0.8970	0.8974	0.8979	0.8980
mse	0.0002	0.0001	0.0001	0.0000	0.0000
bias	-0.0040	-0.0030	-0.0026	-0.0021	-0.0020

From the above tables, one can see that for the different values of q the values of the mse and the bias decrease as the sample size increases which indicates the asymptotic consistency property of the mle. Also for different sample sizes the value of the mse decreases as q increases which means that the estimation of q is more accurate for large values of q. Similarly, the estimation of q is more accurate for large values of α since the mse decreases as α increases. Moreover, all the biases are negative therefore, all the mle's are underestimated. By comparing the three tables

we see that the $DGD(\alpha, q)$ gives more accurate estimation as α increases. Only in the case of the G(q) we see that it is more accurate than the $DGD(2, q)$ for small values of q. In addition, one can note that the $DGD(5, q)$ gives excellent estimates for q even for small sample sizes.

4.1.4 Moment method

The moment estimator of the parameter q of the $DGD(\alpha, q)$ when $\alpha = \alpha_o$ is derived by solving the following equation for q

$$\frac{\sum_{i=0}^{\alpha_o - 1} A(\alpha_o, i) q^i}{(1-q)\sum_{i=0}^{\alpha_o - 2} A(\alpha_o - 1, i) q^i} = \bar{x}\,. \tag{1.15}$$

One can see that solving this equation for q leads to the same estimator given by the maximum likelihood method explained above. Thus, the method of moment estimator and the mle of q coincide when α is known.

On the other hand, the moment estimators of the $DGD(\alpha, q)$ parameters when both α and q are unknown are derived by solving the following equations for α and q

$$\frac{\sum_{i=0}^{\alpha - 1} A(\alpha, i) q^i}{(1-q)\sum_{i=0}^{\alpha - 2} A(\alpha - 1, i) q^i} = \bar{x} \tag{1.16}$$

$$\frac{\sum_{i=0}^{\alpha} A(\alpha + 1, i) q^i}{(1-q)^2\sum_{i=0}^{\alpha - 2} A(\alpha - 1, i) q^i} = s^2\,. \tag{1.17}$$

There is no explicit solution for these two equations, so we need to solve them numerically.

4.2 Estimation of the GDLD(α,q)

4.2.1 Generating sample from the GDLD(α,q)

Samples of the $GDLD(\alpha, q)$ are generated by solving equation (1.1) for t_i. For example, samples of the $GDLD(2, q)$ are generated by solving the following equations for t_i

$$\frac{2 - q + q^{2+t_i} + q^{1+t_i}(t_i - 1) - q^{t_i}(1 + t_i)}{2 - q} - u_i = 0. \tag{2.1}$$

Here, $x_i = [y_i]$, where $y_i = |t_i|$

4.2.2 Maximum likelihood method

The Likelihood function of the GDLD(2,q) for a sample of size n is given by

$$L(q) = \frac{(1-q)^{2n}}{(2-q)^n} q^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (1 + \frac{x_i}{q}). \tag{2.2}$$

Thus, the log likelihood function is

$$\ln L(q) = 2n \ln(1-q) - n \ln(2-q) + \ln q \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \ln(1 + \frac{x_i}{q}). \tag{2.3}$$

Differentiating (2.3) with respect to q and then equating to zero gives the following equation which will be solved numerically for q:

$$-\frac{2n}{1-q} + \frac{n}{2-q} - \frac{1}{q} \sum_{i=1}^{n} \frac{x_i}{(q + x_i)} + \frac{1}{q} \sum_{i=1}^{n} x_i = 0. \tag{2.4}$$

Similarly, the equation that gives the mle of the parameter q of the GDLD(3,q) is given by:

$$-\frac{3n}{1-q}+\frac{2n(1-q)}{(1-q)^2+1}-\frac{(1+2q)}{q(1+q)}\sum_{i=1}^{n}\frac{x_i^2}{q(1+q)+x_i^2}+\frac{1}{q}\sum_{i=1}^{n}x_i=0\,. \tag{2.5}$$

Next, samples of sizes 10, 20, 30, 50 and 100 are generated from the GDLD(2,q) and the GDLD(3,q) using (1.1) by setting the population value of q to be 0.4, 0.5, 0.7 and 0.9 . The mle's of q which are The average estimates taken over 1000 replicates are given in tables 6 and 7 also the mean square error and the bias of the estimates are calculated.

Table 6. mle of q for the GDLD(2,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.2722	0.2629	0.2645	0.2686	0.2711
mse	0.0289	0.0258	0.0233	0.0203	0.0181
bias	-0.1278	-0.1371	-0.1355	-0.1314	-0.1289
(q=0.5)					
$\hat{q}_{mle}$	0.3612	0.3721	0.3784	0.3826	0.3857
mse	0.0323	0.0237	0.0195	0.0167	0.0145
bias	-0.1388	-0.1279	-0.1216	-0.1174	-0.1144
(q=0.7)					
$\hat{q}_{mle}$	0.6344	0.6450	0.6487	0.6524	0.6541
mse	0.0114	0.0062	0.0047	0.0034	0.0027
bias	-0.0657	-0.0550	-0.0513	-0.0476	-0.0459
(q=0.9)					
$\hat{q}_{mle}$	0.8890	0.8920	0.8931	0.8943	0.8948
mse	0.0008	0.0004	0.0002	0.0001	0.0001
bias	-0.0110	-0.0080	-0.0069	-0.0057	-0.0052

Table 7. mle of q for the GDLD(3,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}_{mle}$	0.2876	0.2982	0.3016	0.3044	0.3058
mse	0.0197	0.0144	0.0123	0.0107	0.0097
bias	-0.1124	-0.1018	-0.0984	-0.0956	-0.0942
(q=0.5)					
$\hat{q}_{mle}$	0.4116	0.4211	0.4238	0.4268	0.4283
mse	0.0141	0.0097	0.0081	0.0067	0.0058
bias	-0.0884	-0.0789	-0.0762	-0.0732	-0.0717
(q=0.7)					
$\hat{q}_{mle}$	0.6610	0.6660	0.6680	0.6702	0.6710
mse	0.0047	0.0027	0.0020	0.0015	0.0012
bias	-0.0390	-0.0340	-0.0320	-0.0298	-0.0290
(q=0.9)					
$\hat{q}_{mle}$	0.8931	0.8949	0.8955	0.8964	0.8966
mse	0.0004	0.0002	0.0001	0.0001	0.0000
bias	-0.0066	-0.0051	-0.0045	-0.0036	-0.0034

Tables 6 and 7 show that for the different values of q the values of the mse and the bias decrease as the sample size increases which indicates the asymptotic consistency property of the mle. Also for different sample sizes the value of the mse decreases as q increases which means that the estimation of q is more accurate for large values of q. Similarly, the estimation of q of the GDLD(3,q) is more accurate than that of the GDLD(2,q), since the mse decreases as α increases. Moreover, all the biases are negative therefore, all the mle's are underestimated.

4.2.3 Moment method

The moment estimator of the parameter q of the $GDLD(2,q)$ is derived by solving the following equation for q

$$\frac{q(2-q)+1}{(1-q)(2-q)} = \bar{x}\,. \tag{2.6}$$

And this leads to

$$\hat{q} = \frac{-2-3\bar{x} \pm \sqrt{8+8\bar{x}+\bar{x}^2}}{-2(1+\bar{x})}\,. \tag{2.7}$$

Whereas, for the $GDLD(3,q)$ the moment estimator of the parameter q is derived by solving the following equation numerically for q

$$\frac{p^2 q(1+q)+1+4q+q^2}{(p^2+1)(1+q)p} - \bar{x} = 0\,. \tag{2.8}$$

Next, samples of different sizes are generated from the GDLD(2,q) and the GDLD(3,q) using (1.1) by setting the population value of q to be 0.4, 0.5, 0.7 and 0.9. The moment estimates of q which are The average estimates taken over 1000 replicates are given in table 8 and 9 also the mean square error and the bias of the estimates are calculated.

Table 8. moment estimates of q for GDLD(2,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}$	0.1951	0.1892	0.1905	0.1940	0.1972
mse	0.0559	0.0539	0.0513	0.0474	0.0437
bias	-0.2049	-0.2108	-0.2095	-0.2060	-0.2028

(q=0.5)					
$\hat{q}$	0.3283	0.3411	0.3474	0.3530	0.3563
mse	0.0473	0.0349	0.0296	0.0254	0.0226
bias	-0.1717	-0.1589	-0.1526	-0.1470	-0.1437
(q=0.7)					
$\hat{q}$	0.6294	0.6400	0.6438	0.6478	0.6494
mse	0.0127	0.0070	0.0054	0.0040	0.0032
bias	-0.0706	-0.0600	-0.0562	-0.0522	-0.0506
(q=0.9)					
$\hat{q}$	0.8889	0.8919	0.8930	0.8942	0.8947
mse	0.0008	0.0004	0.0003	0.0001	0.0001
bias	-0.0111	-0.0081	-0.0070	-0.0058	-0.0053

Table 9. moment estimates of q for GDLD(3,q)

	n=10	n=20	n=30	n=50	n=100
(q=0.4)					
$\hat{q}$	0.2738	0.2808	0.2840	0.2874	0.2885
mse	0.0249	0.0189	0.0166	0.0145	0.0134
bias	-0.1262	-0.1192	-0.1160	-0.1126	-0.1115
(q=0.5)					
$\hat{q}$	0.4056	0.4125	0.4149	0.4183	0.4194
mse	0.0165	0.0114	0.0097	0.0081	0.0073
bias	-0.0944	-0.0875	-0.0851	-0.0817	-0.0806
(q=0.7)					
$\hat{q}$	0.6600	0.6647	0.6666	0.6688	0.6695
mse	0.0049	0.0028	0.0022	0.0016	0.0013
bias	-0.0400	-0.0353	-0.0334	-0.0312	-0.0305

(q=0.9)					
$\hat{q}$	0.8931	0.8949	0.8955	0.8964	0.8966
mse	0.0004	0.0002	0.0001	0.0001	0.0000
bias	-0.0069	-0.0051	-0.0045	-0.0036	-0.0034

From tables 8 and 9, one can note that the mse and the bias decrease as the sample size increases. Also, the value of the mse decreases as q increases which means that the moment estimation of q is more accurate for large values of q.

Comparing tables 6 and 7 with 8 and 9, one can see that when q is small (q=0.4), the mle's are more accurate than the moment estimate in terms of mean square error. Whereas, for large values of q (q=0.7 and 0.9), the mle's and the moment estimates are almost the same.

4.3 Estimation of the $GDLD_2(\alpha, q)$

4.3.1 Generating sample from the $GDLD_2(2, q)$

Samples of the $GDLD_2(2, q)$ are generated by solving the following equation for t_i.

$$\frac{-2-q+q^2+t_i q^{t_i+3}+(1-3t_i-2t_i^{\,2})q^{t_i+1}-(1+t_i-t_i^{\,2})q^{t_i+2}+(2+3t_i+t_i^{\,2})q^{t_i}}{q^2-q-2}-u_i=0 \qquad (3.1)$$

Here, $x_i=[y_i]$ *where* $y_i=|t_i|$.

4.3.2 Maximum likelihood method

The Likelihood function of the $GDLD_2(2, q)$ for a sample of size n is given by

$$L(q)=\frac{(1-q)^{3n}}{(2-q)^{n}}q^{\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}\left(\frac{x_i}{q}+\frac{x_i^2}{q(1+q)}\right). \tag{3.2}$$

Thus, the log likelihood function is

$$\ln L(q)=3n\ln(1-q)-n\ln(2-q)+\ln q\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}\ln\left(\frac{x_i}{q}+\frac{x_i^2}{q(1+q)}\right). \tag{3.3}$$

Differentiating (2.3) with respect to q and then equating to zero gives the following equation which will be solved numerically for q:

$$-\frac{3n}{1-q}+\frac{n}{2-q}+\frac{1}{q}\sum_{i=1}^{n}x_i-\frac{1}{q(1+q)}\sum_{i=1}^{n}\left(\frac{(1+q)^2x_i+(1+2q)x_i^2}{(1+q)x_i+x_i^2}\right)=0. \tag{3.4}$$

Next, a simulation study is carried out by generating samples of different sizes from the $GDLD_2(2,q)$ using (1.1). The population values of q are 0.4, 0.8 and 0.9. The mle's of q which are The average estimates taken over 1000 replicates are given in table 10 also the mean square error and the bias of the estimates are calculated.

Table 10. mle of q for $GDLD_2(2,q)$

	n=10	n=20	n=30	n=40	n=50
(q=0.4)					
$\hat{q}_{mle}$	0.30831	0.31364	0.31449	0.31415	0.31870
mse	0.01443	0.01054	0.00936	0.00896	0.00794
bias	-0.09169	-0.08636	-0.08551	-0.08585	-0.08130
(q=0.8)					
$\hat{q}_{mle}$	0.77913	0.77696	0.77711	0.77723	0.79587
mse	0.00187	0.00134	0.00099	0.00089	0.00031
bias	-0.02087	-0.02304	-0.02289	-0.02277	-0.00413

(q=0.9)					
$\hat{q}_{mle}$	0.89281	0.89442	0.89500	0.89520	0.89710
mse	0.00044	0.00022	0.00016	0.00011	0.00009
bias	-0.00719	-0.00558	-0.00410	-0.00479	-0.00290

From the above table one can see that, q is slightly better estimated when the sample size increased, whereas it is much better estimated when the population value of q increased. Furthermore, all the mle's are underestimated.

4.3.3 Moment method

The moment estimator of the parameter q of the $GDLD_2(2,q)$ is derived by solving the following equation for q

$$\frac{2+5q-q^3}{2-q-2q^2+q^3}-\bar{x}=0. \tag{3.5}$$

Next, samples of different sizes are generated from the $GDLD_2(2,q)$ and using (1.1) by setting the population value of q to be 0.4, 0.8 and 0.9. The moment estimates of q which are The average estimates taken over 1000 replicates are given in table 11 also the mean square error and the bias of the estimates are calculated.

Table 11. moment estimates of q for $GDLD_2(2,q)$

	n=10	n=20	n=30	n=40	n=50
(q=0.4)					
$\hat{q}$	0.30847	0.31413	0.31445	0.31549	0.31789
mse	0.01438	0.01042	0.00938	0.00864	0.00804
bias	-0.09153	-0.08587	-0.08555	-0.08451	-0.08211

(q=0.8)					
$\hat{q}$	0.77776	0.78202	0.78267	0.78387	0.78456
mse	0.00216	0.00111	0.00085	0.00063	0.00056
bias	-0.02224	-0.01798	-0.01733	-0.01613	-0.01544
(q=0.9)					
$\hat{q}$	0.89265	0.89478	0.89511	0.89573	0.89606
mse	0.00046	0.00022	0.00016	0.00011	0.00009
bias	-0.00735	-0.00522	-0.00489	-0.00427	-0.00394

Table 11 shows, for the different values of q the values of the mse and the bias decrease as the sample size increases. Also for different sample sizes the value of the mse decreases as q increases which means that the estimation of q is more accurate for large values of q. Moreover, all the biases are negative therefore, the moment estimates are underestimated. Comparing tables 10 and 11 one can see that, according to the mse both methods have almost the same precision of estimating q.

4.4 Real-life data applications

The first data set

The data represents the survival times (in days) of 40 patients suffering from blood cancer. The data is collected in one of the Ministry of Health Hospitals in Saudi Arabia, see Abouammoh et al. (1994).

115, 181, 255, 418, 441, 461, 516, 739, 743, 789, 807, 865, 924, 983, 1024, 1062, 1063, 1165, 1191, 1222, 1222, 1251, 1277, 1290, 1357, 1369, 1408, 1455, 1478, 1549, 1578, 1578, 1599, 1603, 1605, 1696, 1735, 1799, 1815, 1852.

We fit the $DGD(\alpha,q)$, $NB(r,q)$, $GDLD(\alpha,q)$ and $GDLD_2(\alpha,q)$ when α and r are known to this data and the Kolmogorov-Smirnov test presented in chapter 1 is utilized to determine the appropriateness of the model. Our hypotheses are

H_o : The data follow the hypothesized distribution.

H_1 : The data doesn't follow the hypothesized distribution.

The next tables give the maximum likelihood estimates of the parameter q of each distribution when α and r are known. Also for some cases the estimates of the survival function and the FR are calculated at the sample mean $\bar{x}$ using the invariant property of the mle of q, means $\hat{S}=S(\hat{q})$ and $\hat{h}=h(\hat{q})$. The K- S distances between the fitted and the empirical distributions are also shown.

Table 12. Results of fitting the $DGD(\alpha_o,q)$ and $NB(r_o,q)$

	$\hat{q}$	K-S	$\hat{S}$	$\hat{h}$
G(q)	0.99912	0.27824	0.36772	0.00088
DGD(2,q) NB(2,q)	0.99824 0.99824	0.17346 0.17368	0.40577 0.40577	0.00117 0.00117
DGD(3,q) NB(3,q)	0.99737 0.99737	0.14345 0.14342	0.67644 0.42290	0.00070 -0.00002
DGD(4,q) NB(4,q)	0.99649 0.99649	0.13587 0.13584	0.85697 0.43313	0.00037 0.00158

Table 13. Results of fitting the $GDLDI, II$

	$\hat{q}$	K-S
$GDLD(2,q)$	0.99824	0.17368
$GDLD(3,q)$	0.99737	0.14344
$GDLD_2(2,q)$	0.99737	0.14343

From the above tables one can note that the values of the modified K-S statistic are less than the critical value which is 0.21. Hence we can conclude, this data fits all the hypothesized distribution for the given values of α and r. Moreover, the value of the K-S statistic decreases as α and r increase which indicates that the $DGD(\alpha,q)$, $NB(r,q)$ and the $GDLD(\alpha,q)$ fit this data better for large values of α and r.

Figures 18, 19, and 20 show the p-p plots for fitting the $DGD(\alpha,q)$ when $\alpha = 2,3,4$. These figures suggest that the $DGD(\alpha,q)$ is an equitable model for the data and it fits better as α increase since the deviation from straightening decreases in this case.

Based on the values of the K-S statistic, the $DGD(\alpha,q)$ provides a better fit compared to the $NB(r,q)$ when $\alpha = r = 2$. However, they both give approximately the same fit when $\alpha = r = 3$ and $\alpha = r = 4$. Moreover, the mixture $GDLD_2(2,q)$ provide a better fit to the data compared to the other distribution with the same shape parameter 2. It is clear from table 12 that the estimate survival function increases as α and r increase while the estimate FR decreases.

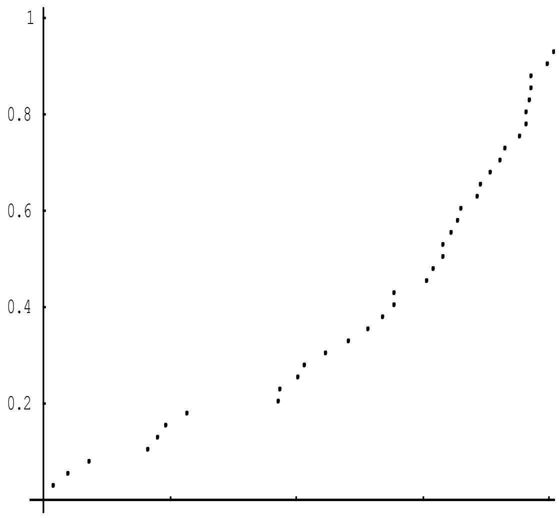


Figure 18. P-P plot for fitting the DGD(2,q)

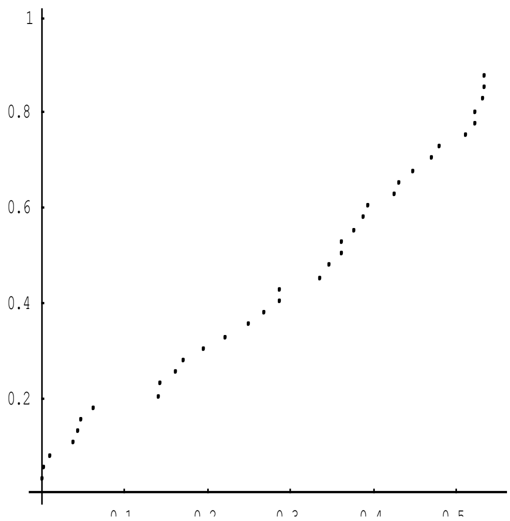


Figure 19. P-P plot for fitting the DGD(3,q)

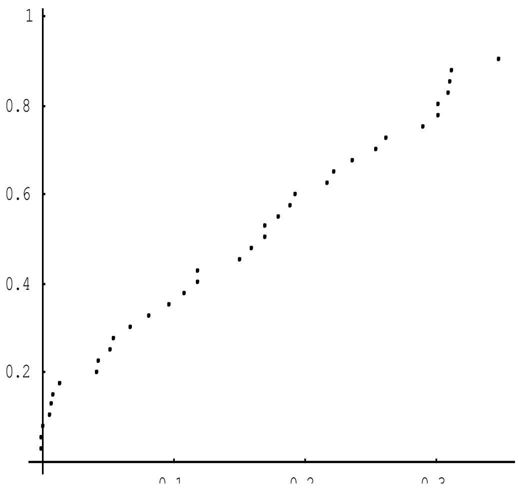


Figure 20. P-P plot for fitting the DGD(4,q)

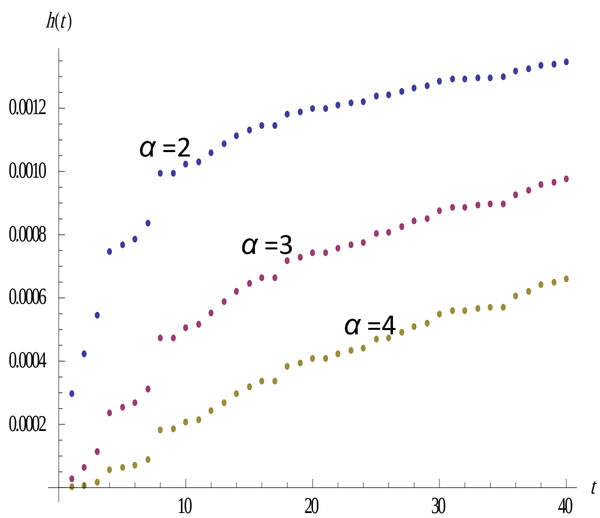


Figure 21. FR for the cancer patients data

The Second data set

The data represents 18 lifetimes of certain electronic devices given in Wang (2000).

5, 11 ,21, 31, 46, 75, 98, 122, 145 ,165 ,196 ,224 ,245, 293, 321, 330, 350 ,420.

We fit the $DGD(\alpha,q)$, $GDLD(\alpha,q)$ and the $GDLD_2(\alpha,q)$ when α is known to this data. The Kolmogorov-Smirnov test is used to determine the appropriateness of the model.

The next table give the mle's of the parameter q of each distribution when α is known. Also The values of the K- S statistic are calculated.

Table 14. Results of fitting some distributions to second data

	$\hat{q}$	K-S
$DGD(2,q)$	0.98845	0.17511
$DGD(3,q)$	0.98272	0.22883
$GDLD(2,q)$	0.98851	0.17249
$GDLD(3,q)$	0.98273	0.22871
$GDLD_2(2,q)$	0.982818	0.227056

From table 14, it is clear that the values of the modified K-S statistic are less than the critical value which is 0.309. Hence we can conclude that this data fits all the given distributions and it fits the $GDLD(2,q)$ most.

References

Abouammoh, A. M. (1988). On the criteria of the mean remaining life. Statist. & prob. Letters, 6, 205-211.

Abouammoh, A. M. and Mashhour, A. F. (1981). A note on the unimodality of discrete distributions. Comm. Statistic, Ser. A, 10, 1345-1354.

Abouammoh, A, M., Abdulghani, S. A. and Qamber, I. S. (1994). On partial orderings and testing of new better than renewal used classes. Reliability Engineering and System Safety, 43, 37-41.

Abouammoh, A. M. and Alhazzani, Najla. S. (2012). On discrete gamma distribution (submitted for publication).

Ali Khan, M. S., Khalique, A. and Abouammoh, A. M. (1989). On estimating parameters in a discrete Weibull distributions. IEEE Trans. Reliability, 38, 347-350.

Barlow, R. E. and Proschan, F. (1981). Statistical Theory of Reliability and Life Testing. To Begin With, Silver Spring, MD.

Bryson, M. C. and Siddiqui, M. M. (1969). Some criteria for aging. J. Amer. Statist. Assoc., 64, 1472-1483.

del Castillo, J. and Perez-Casany, M. (2005). Over-dispersed and under-dispersed Poisson generalizations. Journal of Statistical Planning and Inference, 134, 486-500.

Ghitany, M. E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. Math. Comput. Simulat., 78, 493-506.

Ghitany, M. E. and Almutairi, D. K. (2008). Size-biased Poisson-Lindley distribution and its application. International Journal of Statistics, 3, 299-311.

Gupta, P. L., Gupta, R. C. and Tripathi, R. C. (1997). On the monotone property of discrete failure rate. J. of Statistical Planning and Inference, 65, 255-268.

Hirzebruch, F. (2008). Eulerian polynomials. Munster J. of Math, 1 , 9–14.

Johnson, N. L., Kots, S. and Kemp, A. (1992). Univariate Discrete Distribution. New York, Wiley.

Klugman, S. A., Panjer, H. H. and Willmot, G. E. (2008). Loss Models: From Data to Decision. USA, Wiley.

Lawless, J. F. (2003). Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York.

Law, A. and Kelton, W. (1991). Simulation Modeling and Analysis. McGraw-Hill, Inc. New York.

Leemis, L. M. (2003). Reliability: Probabilistic Models and Statistical Methods. Prentice hall. inc, New Jersey.

Lemaire, J. (1979). How to define a bonus-malus system with an exponential utility function. ASTIN Bull., 10, 274-282.

Lindley, D. V. (1958) Fiducial distributions and Bayes theorem. JRSS, 20, 102-107.

Lord, D. and Geedipally, S. R. (2011). The negative binomial-Lindley distribution as a Tool for analyzing crash data characterized by a large amount of zeros. (Submitted for publication).

Shengwang, M. and Wei, Y. (1999). Accounting for individual over-dispersion in a bonus-malus automobile insurance system. ASTIN Bull., 29, 327-337.

Mir, K. A. (2009). On size-biased negative binomial distribution and its use in zero-truncated cases. Measurement Science Review, Volume 9, Issue 2, 33-35.

Mood, A. M., Graybill, F. A. and Boes, D. C. (1985). Introduction to the Theory of Statistic. 3^{rd} ed. Mc Graw Hill.

Nakagawa, T. and Osaki, S. (1975). The discrete Weibull distributions. IEEE Trans. Reliab., 24, 300-301.

Padgett, W. J. and Spurrier, J. D. (1985). On discrete failure models. IEEE Trans. Reliab., 34, 253–256.

Patil, G. P. and Ord, J. K. (1975). On size-biased sampling and related form-invariant weighted distributions, Sankhya, 38, 48-61.

Patil, G. P. and Rao, C. R. (1977). The weighted distributions: a survey of their applications, in: P. R. Krishnaiah (Ed.), Applications of Statistics, Amsterdam, North-Holland, 383-405.

Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics, 34, 179-189.

Salvia, A. A. and Bollinger, R. C. (1982). On discrete hazard function. IEEE transactions on Reliability, Vol. 31, No. 5, 458-459.

Sankaran, M. (1970). The discrete Poisson-Lindley distribution. Biometrics, 26, 145-149.

Shapiro, S. S. and Brain, C. W. (1981). A review of distributional testing procedures and development of a censored sample distributional test. In Statistical distributions in scientific work,Vol. 5, ed. C. Taillie, G.P. Patil and B.A. Baldessari, 1-24. Dordrecht, Holland: D. Reidel Publishing Company.

Stacy, E. W. (1962). A generalization of the gamma distribution. Ann. Math. Statist., 33, 1187.

Stein, W. E. and Dattero, R. (1984). A new discrete Weibull distribution. IEEE Trans. Reliab., 33, 196-107.

Wang, F. K. (2000). A new model with bathtub-shaped failure rate using an additive Burr XII distribution. Reliability Engineering and System Safety, 70, 305-312.

Zamani, H. and Ismail, N. (2010). Negative binomial–Lindley distribution and its applications. Journal of Mathematics and Statistics, 6 No1, 4–9.