# ON A TWO-PARAMETER DISCRETE DISTRIBUTION AND ITS APPLICATIONS

**Shovan Chowdhury**

Department of Quantitative Methods and Operations Management,

Indian Institute of Management Kozhikode; India

E-mail: shovanc@iimk.ac.in

**Amitava Mukherjee**

Department of Quantitative Methods and Information Systems,

Indian Institute of Management Udaipur; India

and

**Asok K. Nanda**

Department of Mathematics and Statistics,

IISER Kolkata, Mohanpur Campus, West Bengal, India

**Abstract:** Here we introduce two-parameter compounded geometric distributions with monotone failure rates. These distributions are derived by compounding geometric distribution and zero-truncated Poisson distribution. Some statistical and reliability properties of the distributions are investigated. Parameters of the proposed distributions are estimated by the maximum likelihood method as well as through the minimum distance method of estimation. Performance of the estimates by both the methods of estimation are compared based on Monte-Carlo simulations. An illustration with Air Crash casualties demonstrates that the distributions can be considered as a suitable model under several real situations.

**Keywords:** Compounding; Geometric Distribution; Hazard rate function; Maximum likelihood estimation; Method of minimum distance; Monte-Carlo; Zero-Truncated Poisson Distribution.

**AMS 2010 subject classification**: Primary 62E99; Secondary 62N05, 62F10

# 1 Introduction

In theoretical and applied statistics, duration of life of certain organisms, devices, structures, materials etc. is one of the well known random variables that attracts interest of many researchers from various scientific disciplines. In reliability engineering, to assess or to determine the lifespan of a system of components, for example, bearings, seals, gears etc., sample information on failure history are collected in suitable fashion from the population. A major challenge to the statisticians and the reliability engineers is to develop an appropriate model for the observed failure time data. In this context, a substantial part has been devoted to the mathematical description of the length of life by a continuous failure time distribution.

However, in various practical scenarios, discrete failure time distributions are appropriate to model lifetime or the failure rate of the concerned object. For example, a discrete distribution is appropriate when (i) a piece of equipment operates in cycles and the number of cycles prior to failure is observed, (ii) failures occur only due to incoming shocks and number of rounds fired until failure becomes more crucial than age at failure, (iii) a device is monitored only once per time period (e.g., an hour, a day) and the observation is the number of time periods successfully completed prior to failure of the device. When data are collected from a continuous distribution, one will have to round it off. So, to be specific, data of continuous nature may be obtained only in theory, but in practice, data are always discrete in nature. Allison (1982) discussed discrete time models for the analysis of event histories in the context of Sociological methodologies. Hamerle (1986) addressed certain techniques of regression analysis for discrete event history or failure time data. Khalique (1989) studied some inferential aspects of certain class of discrete failure-time distributions. Interested readers may also see Salvia and Bollinger (1982), Padgett and Spurrier (1985), Adams and Watson (1989), Klar (1999) and the references therein.

Fahrmeir (1994) considered some dynamic modeling and penalized likelihood estimation methods for discrete time survival data, whereas Fahrmeir and Knorr-Held (1997) proposed estimation of dynamic discrete-time duration models via Markov Chain Monte Carlo. Rocha-Martine and Shaked (1995) proposed a discrete-time model of failure and repair. Fahrmeir and Wagenpfeil (1996) investigated time-varying effects in discrete duration for competing risk models. Scheike and Jensen (1997) proposed an interesting discrete survival model with random effects in connection with time to pregnancy. Shis (1998) considered multivariate discrete failure time data.

Klar (1999) nicely summarised theory and methods for model selection for a family of discrete failure time distributions. He pointed out that though a large number of new discrete distributions have appeared in the literature, there is dearth of flexible discrete models which, at the same time, allow for easy statistical inference even though exhibit constant, increasing or decreasing failure rates. The scenario has changed only a little even after a decade though the importance of discrete lifetime distributions has well recognized among researchers and practitioners in Sociology and in Biometrics, among several other practical fields. It is well known that this is one of the major areas of interest to the reliability engineers.

In the recent times, host of researchers investigated discrete failure time models. Jiang and Kumar (2001) studied failure diagnosis of discrete event systems with linear-time temporal logic fault specifications. Belzunce et al. (2009) discussed ageing properties of a discrete time failure and repair model. Singh et al. (2009) analyzed the dynamic system model with discrete failure time distribution. Patil and Bagkavos (2012) proposed semiparametric smoothing of discrete failure time data.

In the present article, we propose two discrete distributions arising from the following motivating example:

Suppose a financial institution (FI) has $m$ (known positive integer) regional offices. Further assume that $i$th regional office (RO) controls functioning of $N_i$ $(i = 1, 2, ..., m)$ branch offices (BO) which work independently of each other. One of the indicators for the evaluation of the performance of each of the BO's under their respective RO's is the number of new accounts opened in the last three months each exceeding a high amount, $\tau$, say, which is known. Let us call these accounts as High Deposit Accounts (HDA). Quarterly reviews on the performance of the BO's across all the RO's may incur some extra operational costs that may not be worthy to spend. A probabilistic model for the performance review may be recommended henceforth.

Let $M_{ij}$ be the number of new HDA's opened in the last three months at the $j$th branch of the $i$th regional office. Then, $M_{i1}, M_{i2}, ..., M_{iN_i}$ are assumed to be independent and identical geometric random variables. We assume that $N_i$ follows zero-truncated Poisson distribution independent of $M_{ij}$. If $M_i$ denotes the minimum of the number of HDA's at the $i$th RO, we can write $M_i = \min_{1 \le j \le N_i} M_{ij}$ and may be termed as In-Bound Performance Index ($IBPI_i$) of the $i$th RO.

With similar arguments, the FI may be interested to know the number of old HDA's closed in the last three months each of which is less than a given amount, say, $\phi$ which is known. Let $M^*_{ij}$ be the number of old HDA's closed in the last three months at the $j$th branch of the $i$th regional office. If $M^*_i$ denotes the maximum of the number of HDA's at the $i$th RO, we can write $M^*_i = \max_{1 \le j \le N_i} M^*_{ij}$ and may be termed as Out-Bound Performance Index ($OBPI_i$) of the $i$th RO.

In the current context, we study the distributions of $IBPI_i$ and $OBPI_i$ which are described as Compounded Geometric (CG) distribution where the distribution of the observed discrete data is described by a geometric distribution, and a zero-truncated Poisson distribution is used for the distribution of sample size of each subgroup. We develop two types of CG distributions - one, from the minimum (e.g. the case of $IBPI_i$), and another from the maximum (e.g. the case of $OBPI_i$) of the observed data from each subgroup.

The rest of the paper is organized as follows. In Section 2, a new distribution is obtained by mixing the geometric distribution and a zero-truncated Poisson distribution. In Section 3, various properties of the introduced distribution are discussed. Parameters of the distribution are estimated in Section 4 by the maximum likelihood method and the distance minimization method. In Section 5, a simulation study is performed to show the behaviour of asymptotic variances and covariances of maximum likelihood estimators (MLEs). An illustrative example based on real life data is provided in Section 6. Finally, Section 7 concludes the manuscript with a future research problem. It is to mention that by $\theta_1 \stackrel{sign}{=} \theta_2$ we mean that $\theta_1$ and $\theta_2$ have the same sign. In the manuscript, some of the proofs are trivial and hence are given in the Appendix.

## 2 The Compounded Geometric distribution

The geometric distribution with parameter $p \in (0,1)$, denoted as $G(p)$, has the probability mass function (pmf) $g_K$ given by

$$g_K(k;p) = pq^k,\ \ k = 0,1,2,...;\ \ p \in (0,1), q = 1-p, \tag{2.1}$$

with cumulative distribution function (CDF) given by

$$G_K(k;p) = 1 - q^{k+1}. \tag{2.2}$$

Let $K_1, K_2, ..., K_M$ be a random sample from $G(p)$ and $M$ be a zero-truncated Poisson variable with pmf given by

$$P(m;\lambda) = \frac{e^{-\lambda}\lambda^m(1-e^{-\lambda})^{-1}}{\Gamma(m+1)},\ \ m \in \mathcal{N}, \lambda > 0, \tag{2.3}$$

where $\mathcal{N}$ is the set of natural numbers. Assuming that random variables $K$ and $M$ are independent, we define

$$U = \min_{1 \le i \le M} K_i.$$

Then the conditional distribution of $(U \mid M = m)$

$$F_U(u \mid m) = 1 - (1 - G_K(u;p))^m = 1 - q^{(u+1)m}, \tag{2.4}$$

with the corresponding pmf

$$f_U(u \mid m) = P(U \le u \mid m) - P(U \le u-1 \mid m) = q^{um}(1-q^m). \tag{2.5}$$

Thus, the unconditional pmf of $U$ is given by

$$f(u;q,\lambda) = \sum_{m=1}^{\infty} f_U(u \mid m)P(m;\lambda) = \frac{e^{\lambda q^u} - e^{\lambda q^{u+1}}}{e^{\lambda}-1},\ \ u = 0,1,2,...;\ \ q \in (0,1), \lambda > 0. \tag{2.6}$$

Hereafter, the distribution of $U$ will be referred to as the CG distribution of Type-I and will be denoted by CG(I). As $\lambda$ approaches zero, the CG(I) distribution leads to the geometric distribution. Sometimes, to be specific, we write CG(I)(p, $\lambda$) to mean that the Type-I CG has been derived based on a CG(I) distribution with parameters $p$ and $\lambda$.

In this context, it may be interesting to note that, writing

$$V = \max_{1 \le i \le M} K_i,$$

we may get another kind of CG distribution, say CG distribution of Type-II and may be denoted by CG(II). Using Equations (2.1)-(2.3), the conditional distribution of $(V \mid M = m)$ is given by

$$F_V(v \mid m) = (G_K(v;p))^m = \left(1 - q^{v+1}\right)^m, \tag{2.7}$$

with the conditional pmf obtained as

$$f_V(v \mid m) = P(V \leq v \mid m) - P(V \leq v-1 \mid m) = (1 - q^{v+1})^m - (1 - q^v)^m. \tag{2.8}$$

Thus, the unconditional pmf of $V$ is derived as

$$f^*(v; q, \lambda) = \frac{e^{\lambda(1-q^{(v+1)})} - e^{\lambda(1-q^v)}}{e^\lambda - 1},\ v = 0, 1, 2, ...;\ q \in (0,1), \lambda > 0. \tag{2.9}$$

As $\lambda$ approaches zero, CG(II) also leads to its parent geometric distribution. Sometimes we write CG(II)(p,$\lambda$) to mean that the Type-II CG has been derived based on a CG(II) distribution with parameters $p$ and $\lambda$. The following easy-to-show theorem talks of the shape of the CG(I) distribution.

**Theorem 2.1.** *Irrespective of the choices of parameters, the pmf of CG(I) distribution is strictly decreasing.*

**Remark 2.1.** The calculation of mode of CG(II) distribution is not very straightforward. Thus, we use the following indirect approach to get the mode(s). Denoting the nonnegative half of the real line by $\Re_+$, let us write $f_1(w) = e^{-\lambda q^{w+1}} - e^{-\lambda q^w}, \lambda > 0, q > 0$. Let us, for the time being, pretend that $f_1(w)$ is continuous in $w \geq 0$ and is sufficiently differentiable. Now, $f_1'(w) = 0$ gives $w = \frac{1}{\ln q} \ln \left[\frac{\ln(1/q)}{p\lambda}\right] = w_0$, say. Note that, $p\lambda > 0$ and $\ln(1/q) > 0$, imply that $\left[\frac{\ln(1/q)}{p\lambda}\right] > 0$. As $\ln q < 0$ and $w > 0$, $\ln(1/q) \leq p\lambda \Rightarrow q \geq e^{-p\lambda}$. It can be shown that $f(w_0) \geq f(w)$ for all $w \geq 0$. Hence, if $w_0$ in an integer, then $w_0$ will be the mode; otherwise, either $[w_0]$ or $[w_0]+1$ will be the mode depending on whether $f([w_0]) > f([w_0]+1)$ or $f([w_0]) < f([w_0]+1)$, where, $[\cdot]$ denotes the largest integer function. In case $f([w_0]) = f([w_0]+1)$, the distribution will be bimodal with modes at $w_0$ and $w_0+1$. Hence, CG(II) has mode(s) for $q \geq e^{-p\lambda}$. It is to be noted that the distribution is decreasing and hence its mode exists at $v = 0$ when $q = e^{-p\lambda}$. □

The shape of the pmfs for two types of CG distributions are illustrated for selected values of the parameters in Figurs 2.A and 2.B respectively. It is to be noted that CG(I) degenerates at $u = 0$ for large $\lambda$ and fixed $p$, or for fixed $\lambda$ and large $p$. It is easy to see that

CG(II) also degenerates at $v = 0$ for fixed $\lambda$ and large $p$. Nevertheless, if $p$ is fixed, as $\lambda \to \infty$ right tail probability of CG(II) increases sharply.

<Figures 2.A and 2.B. HERE.>

# 3 Properties of the distribution

In this section we study different properties of the CG distributions viz. distribution function, moments, percentiles including median and quartiles, measures of skewness etc. The hazard rates of the CG distributions are also studied along with different results on stochastic orders. Some of the proofs are listed in the appendix.

## 3.1 Distribution function and moments

The CDF of $U$ is given by

$$H_U(u; q, \lambda) = \begin{cases} 0 & \text{if } u < 0 \\ \frac{e^{\lambda} - e^{\lambda q^{[u]+1}}}{e^{\lambda} - 1} & \text{if } u \geq 0 \end{cases}$$

Similarly, the CDF of $V$ is given by

$$H_V(v; q, \lambda) = \begin{cases} 0 & \text{if } v < 0 \\ \frac{e^{-\lambda q^{[v]+1}} - e^{-\lambda}}{1 - e^{-\lambda}} & \text{if } v \geq 0 \end{cases}$$

Below we see that the mean residual life (MRL) function of CG distributions does not exist. The proof is given in the appendix.

**Result 3.1.** *For strictly positive finite $\lambda$ and $q \in (0, 1)$, the expectations and the higher order moments of neither of the two CG distributions exist. Hence MRL function does not exist.*

## 3.2 Percentiles, Median, Inter-Quartile Range and Skewness

It is easy to see that the $100\xi^{th}(\xi \in [0, 1])$ percentile, say $u_\xi$, of CG(I) may be obtained by solving the equation

$$e^{\lambda} - e^{\lambda q^{[u]+1}} = \xi(e^{\lambda} - 1). \tag{3.1}$$

In other words, $100\xi^{th}$ percentile of CG(I) is

$$u_\xi = \frac{1}{\ln(q)} \ln\left(\frac{1}{\lambda}\ln((1-\xi)e^{\lambda}+\xi)\right) - 1. \tag{3.2}$$

As a consequence, we get the following result. For the proof of the result, please see the Appendix given at the end of the manuscript.

**Result 3.2.1.** *Median ($\tilde{\mu}$), Inter-Quartile Range ($IQR$) and Bowley's coefficient of skewness ($b_{1X}$) of CG(I) are respectively given by*

$$\tilde{\mu} = \frac{1}{\ln(q)} \ln\left(\frac{1}{\lambda}\ln\left(\frac{1}{2}e^{\lambda}+\frac{1}{2}\right)\right) - 1,$$

$$IQR = \frac{1}{\ln(q)} \ln\left(\frac{\ln\left(\frac{1}{4}e^{\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{\lambda}+\frac{1}{4}\right)}\right),$$

*and*

$$b_{1X} = \frac{\ln\left(\frac{\ln\left(\frac{1}{4}e^{\lambda}+\frac{3}{4}\right).\ln\left(\frac{3}{4}e^{\lambda}+\frac{1}{4}\right)}{\left(\ln\left(\frac{1}{2}e^{\lambda}+\frac{1}{2}\right)\right)^2}\right)}{\ln\left(\frac{\ln\left(\frac{1}{4}e^{\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{\lambda}+\frac{1}{4}\right)}\right)}.$$

Similarly, $100\xi^{th}(\xi \in [0,1])$ percentile, say $v_\xi$, of CG(II) may be obtained by solving the equation

$$e^{-\lambda q^{[v]+1}} - e^{-\lambda} = \xi(1-e^{-\lambda}), \tag{3.3}$$

which results in

$$v_\xi = \frac{1}{\ln(q)} \ln\left(-\frac{1}{\lambda}\ln((1-\xi)e^{-\lambda}+\xi)\right) - 1 \tag{3.4}$$

The expressions for median, IQR and measure of skewness are given below. The proof may be obtained in the Appendix.

**Result 3.2.2.** *Median ($\tilde{\mu}$), $IQR$ and Bowley's coefficient of skewness ($b_{1X}$) of CG(II) are respectively given by*

$$\tilde{\mu} = \frac{1}{\ln(q)} \ln\left(-\frac{1}{\lambda}\ln\left(\frac{1}{2}e^{-\lambda}+\frac{1}{2}\right)\right) - 1,$$

$$IQR = \frac{1}{\ln(q)} \ln\left(\frac{\ln\left(\frac{1}{4}e^{-\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{-\lambda}+\frac{1}{4}\right)}\right),$$

*and*

$$b_{1X} = \frac{\ln\left(\frac{\ln\left(\frac{1}{4}e^{-\lambda}+\frac{3}{4}\right).\ln\left(\frac{3}{4}e^{-\lambda}+\frac{1}{4}\right)}{\left(\ln\left(\frac{1}{2}e^{-\lambda}+\frac{1}{2}\right)\right)^2}\right)}{\ln\left(\frac{\ln\left(\frac{1}{4}e^{-\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{-\lambda}+\frac{1}{4}\right)}\right)}.$$

It is interesting to note that skewness of none of CG(I) and CG(II) depend on $q$. Moreover, expressions of $b_{1X}$ for both CG(I) and CG(II) show that the distributions are positively skewed.

## 3.3 Survival and Hazard Rate functions

Survival functions of CG(I) and CG(II) are respectively given by

$$S(l;q,\lambda) = P(U \geq l) = \frac{e^{\lambda q^l} - 1}{e^{\lambda} - 1} \tag{3.5}$$

and

$$S^*(l;q,\lambda) = P(V \geq l) = \frac{1 - e^{-\lambda q^l}}{1 - e^{-\lambda}}. \tag{3.6}$$

The hazard rate of CG(I) is given by

$$h(l;q,\lambda) = \frac{P(U = l)}{P(U \geq l)} = 1 - \frac{e^{\lambda q^{l+1}} - 1}{e^{\lambda q^l} - 1}. \tag{3.7}$$

Similarly, the hazard rate of CG(II) is given by

$$h^*(l;q,\lambda) = \frac{P(V = l)}{P(V \geq l)} = 1 - \frac{1 - e^{-\lambda q^{l+1}}}{1 - e^{-\lambda q^l}}. \tag{3.8}$$

The shape of the hazard rate functions (hrf) of CG(I) and CG(II) are illustrated in Figs. 3.3.A and 3.3.B respectively.

The following theorems show the monotonicity of the failure rate functions of the CG distributions. The proofs are given in the Appendix.

**Theorem 3.3.1.** *CG(I) has decreasing failure rate (DFR).*

**Theorem 3.3.2.** *CG(II) has increasing failure rate (IFR).*

**Remark 3.1.** It can be shown that

(i) $\lim_{l\to\infty} h(l) = p = \lim_{l\to\infty} h^*(l)$. Moreover, $h(0) > p$ and $h^*(0) < p$ give that the hazard rate of CG(I) family of distributions is bounded below by $p$ and that of CG(II) is bounded above by $p$.

(ii) $\lim_{\lambda\to 0} h(l) = p$ and $\lim_{\lambda\to\infty} h(l) = 1$

(iii) $\lim_{\lambda\to 0} h^*(l) = p$ and $\lim_{\lambda\to\infty} h^*(l) = 0$

(iv) $\lim_{p\to 0} h(l) = 0 = \lim_{p\to 0} h^*(l)$ and $\lim_{p\to 1} h(l) = 1 = \lim_{p\to 1} h^*(l)$

<GRAPHS 3.3.A and 3.3.B HERE.>

## 3.4 Results on Ordering

To study different properties and usefulness of various stochastic orders for the proposed model, we write $X_{k:n}$ to denote the $kth$ order statistic of a sample $X_1, X_2, ..., X_n$ of size $n$. We also use the standard notations, $X \leq^{ST} Y, X \leq^{DISP} Y, X \leq^{C} Y, X \leq^{SU} Y$ and $X \leq^{*} Y$ to mean that the random variable $X$ is smaller than the random variable $Y$ in stochastic order, dispersive order, convex transform order, super-additive order and star order respectively. For more information on stochastic orders used below, see, for instance, Shaked and Shanthikumar (2007). Let us consider $U_1$ and $U_2$ to be the random variables following CG(I) distribution with CDFs $H_1$ and $H_2$, pmfs $f_1$ and $f_2$, common scale parameter $\lambda$ and shape parameters $p_1$ and $p_2$ respectively. Then the following theorems hold.

**Theorem 3.4.A.** *For $p_1 \geq p_2$ and fixed $\lambda\ (> 0)$, we have*

(i) $U_1 \leq^{ST} U_2$.

(ii) $U_1 \leq^{DISP} U_2$.

(iii) $U_1 \leq^{C} U_2$.

**Proof.** The proof of (i) is trivial and hence is omitted. To prove (ii), note that $U_1 \leq^{DISP} U_2$ if $H_2^{-1}H_1(u) - u$ is increasing in $u$. A simple calculation shows that $H_2^{-1}H_1(u) - u = (u+1)\frac{\ln(q_1/q_2)}{\ln q_2}$, which is increasing in $u$ if $q_1 \leq q_2$, or equivalently $p_1 \geq p_2$. The proof of (iii) follows from the fact that $U_1 \leq^{C} U_2$ if $H_2^{-1}H_1(u)$ is increasing in $u$, which follows from (ii) if $p_1 \geq p_2$. □

The proof of the following corollary follows from the fact that star order and super-additive order follows from convex transform order as discussed in Theorem 3.4.A (iii) above where $U_1 \leq^{*} U_2$ if $H_2^{-1}H_1(u)/u$ is increasing in $u$ and $U_1 \leq^{SU} U_2$ if $H_2^{-1}H_1(u_1+u_2) \geq H_2^{-1}H_1(u_1) + H_2^{-1}H_1(u_2)$ for all $u_1, u_2$.

**Corollary:** *For $p_1 \geq p_2$ and fixed $\lambda\ (> 0)$, $U_1 \leq^{*} U_2$ and $U_1 \leq^{SU} U_2$.* □

Let $\{X_n\}$ be a sequence of iid nonnegative $CG(I)(p_1, \lambda)$ random variables and $\{Y_n\}$ be another sequence of iid nonnegative $CG(I)(p_2, \lambda)$ random variables. Assume that $X_i$ and $Y_i$ are independently distributed, for $i = 1, 2, \ldots$. Further let $M$ be a zero-truncated Poisson random variable independent of $X_i$ and $Y_i$. Then we have the following theorem.

**Theorem 3.4.B**

(i) *For* $p_1 \geq p_2, X_i \leq^{ST} Y_i \Rightarrow X_{1:M} \leq^{ST} Y_{1:M}$ *and* $X_{M:M} \leq^{ST} Y_{M:M}$ *(Shaked and Wong, 1997).*

(ii) *For* $p_1 \geq p_2, X_i \leq^{DISP} Y_i \Rightarrow X_{1:M} \leq^{DISP} Y_{1:M}$ *and* $X_{M:M} \leq^{DISP} Y_{M:M}$ *(Shaked and Wong, 1997)*

iii *For* $p_1 \geq p_2, X_i \leq^{C} Y_i \Rightarrow X_{1:M} \leq^{C} Y_{1:M} \Rightarrow X_{1:M} \leq^{*} Y_{1:M} \Rightarrow X_{1:M} \leq^{SU} Y_{1:M}$ *and* $X_{M:M} \leq^{C} Y_{M:M} \Rightarrow X_{M:M} \leq^{*} Y_{M:M} \Rightarrow X_{M:M} \leq^{SU} Y_{M:M}$ *(Bartoszewicz, 2001).*

**Remark:** *Results similar to Theorem 3.4.A and its Corollary as well as Theorem 3.4.B hold true for CG(II) with* $p_1 \geq p_2$.

# 4 Estimation of the parameters

As moments of none of the CG(I) and CG(II) exist, one of the most popular methods of parameter estimation, viz. method of moment estimation cannot be employed in the present context. Other than the popular and widely used parameter estimation techniques viz. maximum likelihood (ML) estimation, minimum distance (MD) estimation approach has also been recommended by several authors in varying context of method of estimation. In this section, we consider ML estimation and MD estimation based on Cramér von-Mises criterion using the notion of empirical distribution function (EDF).

## 4.1 Method of Maximum Likelihood Estimation

In this section we derive the maximum likelihood (ML) estimators of the unknown parameters $q(= 1 - p)$ and $\lambda$. Let $U_1, U_2, ..., U_n$ and $V_1, V_2, ..., V_n$ be a random sample of size $n$ drawn from CG(I) and CG(II) distributions respectively with parameters $q(= 1 - p)$ and $\lambda$. Then the log-likelihood functions, $L_1(q,\lambda)$ for CG(I) and $L_2(q,\lambda)$ for CG(II), can respectively be written as

$$L_1(q,\lambda) = -n\log(e^{\lambda}-1) + \sum_{i=1}^{n} \log\left(e^{\lambda q^{u_i}} - e^{\lambda q^{u_i+1}}\right) \tag{4.1}$$

$$L_2(q,\lambda) = -n\log(1-e^{-\lambda}) + \sum_{i=1}^{n} \log\left(e^{-\lambda q^{v_i+1}} - e^{\lambda q^{v_i}}\right) \tag{4.2}$$

The ML Estimates (MLE), if exist, can be obtained by solving the system of equations $\frac{\partial L_1}{\partial q} = 0$ and $\frac{\partial L_1}{\partial \lambda} = 0$ for CG(I) and $\frac{\delta L_2}{\delta q} = 0$ and $\frac{\delta L_2}{\delta \lambda} = 0$ for CG(II) as shown below.

$$\frac{\partial L_1}{\partial q} = \sum_{i=1}^{n} q^{u_i-1}\left(\frac{u_i e^{\lambda q^{u_i}} - (u_i+1)q e^{\lambda q^{u_i+1}}}{e^{\lambda q^{u_i}} - e^{\lambda q^{u_i+1}}}\right) \tag{4.3}$$

$$\frac{\partial L_1}{\partial \lambda} = \frac{-ne^{\lambda}}{e^{\lambda}-1} + \sum_{i=1}^{n} q^{u_i}\left(\frac{e^{\lambda q^{u_i}} - q e^{\lambda q^{u_i+1}}}{e^{\lambda q^{u_i}} - e^{\lambda q^{u_i+1}}}\right) \tag{4.4}$$

$$\frac{\partial L_2}{\partial q} = \sum_{i=1}^{n} q^{v_i-1}\left(\frac{v_i e^{-\lambda q^{v_i}} - (v_i+1)q e^{-\lambda q^{v_i+1}}}{e^{-\lambda q^{v_i+1}} - e^{-\lambda q^{v_i}}}\right) \tag{4.5}$$

$$\frac{\partial L_2}{\partial \lambda} = \frac{-ne^{-\lambda}}{1-e^{-\lambda}} + \sum_{i=1}^{n} q^{v_i}\left(\frac{e^{-\lambda q^{v_i}} - q e^{-\lambda q^{v_i+1}}}{e^{-\lambda q^{v_i+1}} - e^{-\lambda q^{v_i}}}\right) \tag{4.6}$$

It is clear that the explicit form of second order derivatives and consequently that of Hessian Matrices (as well as Fisher Information) for both CG(I) and CG(II) are complicated and are not of much practical interest. Computationally, it is easy to see that the Hessian matrices of both the distributions are nonnegative definite and hence MLE's exist. These estimators can be easily obtained by using the function *optim* from the statistical software $R$ (version R.2.14.1, R Development Core Team, 2011). Since full Hessian matrices are difficult to compute in practice, in the current context, we recommend to use quasi-Newton algorithms, namely the BFGS (named after Broyden, Fletcher, Goldfarb and Shanno) algorithm for numerical maximization of log-likelihood functions. Details of the computational results are provided in subsequent section.

## 4.2 Method of minimum distance

Minimum distance (MD) estimation is a statistical method for fitting a mathematical model to data. Let $X_1, \ldots, X_n$ be a random sample from a population with CDF $F(.;\theta)\colon \theta \in \Theta$, the parameter space and $\Theta \subseteq \Re^k (k \geq 1)$. Further suppose that $F_n(.)$ is the empirical distribution function (EDF) based on the sample and $\hat{\theta}$ is an estimator of $\theta$. Then it is legitimate to assume that $F(x; \hat{\theta})$ is an estimator for $F(x; \theta)$. Let $d[\cdot, \cdot]$ be a functional or the criterion function returning some measure of "distance" between the two arguments. Now, in such a set up, if there exists a $\hat{\theta} \in \Theta$ such that

$$d[F(x; \hat{\theta}), F_n(x)] = \inf_{\theta \in \Theta} d[F(x; \theta), F_n(x)],$$

then $\hat{\theta}$ is referred to as the "minimum distance estimator (MDE)" of $\theta$. Interested readers may see Wolfowitz (1957), Blyth (1970), Drossos and Philippou (1980), Parr and Schucany (1980) and Boos (1982) and Weber et al. (2006). In the recent times, Hanselmann et al. (2007) also considered some parametric density estimation problems by minimizing an analytic distance measure.

In this section, we consider Cramér von-Mises type distance measure for estimating the parameters of CG(I) and CG(II) distributions. For CG(I) distribution, estimates of $\lambda$ and $p$ (or $q = 1 - p$) are obtained by minimizing

$$\sum_{i=1}^{n} \left[ \frac{e^{\lambda} - e^{\lambda q^{(u_i+1)}}}{e^{\lambda} - 1} - F_n(u_i) \right]^2,$$

whereas for CG(II) distribution we minimize

$$\sum_{i=1}^{n} \left[ \frac{e^{-\lambda q^{(v_i+1)}} - e^{-\lambda}}{1 - e^{-\lambda}} - F_n(v_i) \right]^2.$$

The technique is generally easy to program, as it neither requires evaluation of likelihood nor calculation of derivatives. The method is computationally faster compared to ML estimation. Performance of the MD method is found to be almost as good as ML method in most of the situations and is better in certain cases.

# 5 Simulation study

It is not feasible to solve the equations $\frac{\partial L_i}{\delta q} = 0$ and $\frac{\partial L_i}{\delta \lambda} = 0$ $(i = 1, 2)$ explicitly in order to get ML estimates for CG(I) and CG(II) distributions. However, one can easily find the numerical solution applying some suitable optimization techniques. In the present context, we use the in-built *optim* function in R.2.14.1 software for numerical minimization of negative of log-likelihood function. We adopt L-BFGS-B method as in Byrd et al.(1995) which is nothing but a limited-memory modification of the BFGS quasi-Newton method mentioned in subsection 4.1. This method allows box constraints, i.e., each parameter may be given a lower and/or upper bound. The initial values are supposed to satisfy the box constraints. For further details, one may see R documentation. We carry out detailed simulation studies to capture the means and the standard deviations (SD) of the ML estimators of $\lambda$ and $q$. Five thousand replicates of Monte-Carlo experiments are considered in the present investigation. Similar method is applied to find out the MDEs and the means and the SDs of MD estimators.

In order to study the convergence of the estimators to its true value as sample size ($n$) increases, a random sample of size 10, 25, 50, 100, 500 and 1000 are obtained respectively from CG(I) and CG(II) for four choices of $q$ viz. 0.97, 0.95, 0.93 and 0.91 and two choices of $\lambda$ viz. 3 and 5. The means and the SDs of estimators $\hat{\lambda}$ and $\hat{q}$ are obtained by two different techniques viz. ML method and MD method. The results for CG(I) distribution are reported in Table 5.1 and the same for CG(II) are presented in Table 5.2.

It is easy to see from Table 5.1 that, for CG(I) distribution, $\hat{q}$ sharply converges to $q$ in almost all cases, as $n$ increases, irrespective of the choice of method of estimation. For $\lambda = 3$, convergence is faster with MD method when true value of $q$ is 0.91 or 0.93, and elsewhere ML estimate of $q$ converges at a faster rate. Interestingly, for small $n$, in almost all cases, bias of the estimator of $q$ is less if MD method is adopted except when the true value of $\lambda$ is 5 and that of $q$ is 0.91. However, the MD estimator of $\lambda$ in such cases returns a larger bias compared to the corresponding ML estimator. For $\lambda$, MD estimates are better than or as good as the corresponding ML estimates when $n$ is large and $q$ is 0.91 or 0.93. Estimates of $\lambda$ for CG(I) distribution are fluctuating a little for small $n$, but slowly converges to its true value as $n$ increases.

From Table 5.2, we see that for both the methods of estimation, $\hat{\lambda}$ and $\hat{q}$ sharply converge to the true value of $\lambda$ and $q$ respectively for CG(II) distribution. We further see that MD method is unquestionably superior to ML approach for small to moderate $n$. Here superiority is judged on the basis of average bias. The lower is the bias, the better is the approach. For large $n$, however, both the approaches are nearly equally effective except for small sampling fluctuations.

INSERT TABLES 5.1 and 5.2 HERE.

# 6 Applications

In this section, we fit both CG(I) and CG(II) distributions to a real data set taken from *www.planecrashinfo.com*. The aviation accident database of this site includes all civil and commercial aviation accidents of noteworthy interest. The data consist of number of fatalities on board (passengers / crew) and persons killed in ground, if any, in different plane crashes in the calendar year 2011. This comprises of 44 observations on number of casualties in 44 different plane crashes. The data set is given below:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 77 | 9 | 6 | 14 | 6 | 23 | 32 | 18 | 7 | 27 |
| 22 | 10 | 47 | 9 | 85 | 7 | 16 | 80 | 2 | 8 | 38 |
| 11 | 12 | 4 | 21 | 8 | 44 | 30 | 3 | 2 | 19 | 18 |
| 2 | 28 | 8 | 1 | 5 | 8 | 1 | 3 | 5 | 4 | 3 |

For each distribution, we derive the estimates obtained by ML method and by MD method along with Kolmogorov-Smirnov (K-S) statistic for goodness-of-fit test and the corresponding p-value. In order to handle ties in the data set, simulated p-value is used for the K-S test. Under ML approach, standard error (SE) of the estimators, Akaike information criterion (AIC), Bayesian information criterion (BIC) are also obtained. The obtained results are presented in Table 6.1. The results show that the p-values of both the models are reasonably high under both ML and MD methods ensuring suitability of both the distributions for the current data set. These results are further validated by the comparison between observed and expected frequencies in Table 6.2. where expected frequencies for both the models commensurate well with the observed frequencies.

A mechanical search (using 1000x1000 possible combinations of the two parameters) for global maximum likelihood estimates correct up to 3 decimal places returns $\hat{\lambda} = 1.151$ and $\hat{q} = 0.961$, which justifies that L-BFGS-B algorithm performs almost accurately in the case of ML estimation of CG(I). Similar search algorithm for MD estimation in case of CG(I) gives $\hat{\lambda} = 9.342$ and $\hat{q} = 0.993$, and shows that MD approach may lead to multiple solutions based on starting values depending on local optima. For simplicity, we have considered ML estimates as the starting values in the MD approach. Unfortunately, L-BFGS-B algorithm does not converge if the start up value is chosen as $\hat{\lambda} = 9.342$ and $\hat{q} = 0.993$. Interestingly, the choice of $\hat{\lambda} = 9.342$ and $\hat{q} = 0.993$ reduces the p-value of K-S test to 0.3969 but improves the same of chi-squared test to 0.8152. That gives class frequencies as 20.65, 10.42, 5.48, 2.99, 1.70 and 2.76. However, two different sets of solutions do not change the nature of fit drastically. A better algorithm for MDE could be an interesting future research problem.

In a similar manner, a mechanical search for global maximum likelihood estimates of CG(II) returns $\hat{\lambda} = 0.001$ and $\hat{q} = 0.947$ which again establishes accuracy of L-BFGS-B algorithm in case of ML estimation of CG(II). However, for the present data, estimate of $\hat{\lambda}$ hits the lower bound of L-BFGS-B algorithm in case of CG(II) model. Ordinary BFGS gives estimates of $\hat{\lambda}$ and $\hat{q}$ as $(0.000002, 0.9470)$ with the present data set which indicates that the true value of $\lambda$ may tend to 0, which, in turn, implies that the CG(II) converges geometric distribution. This phenomenon is also observed from the observed and the expected frequencies in Table 6.2. The mechanical search algorithm yields $\hat{\lambda} = 1.68$ and $\hat{q} = 0.904$ which gives unusually poor fit to CG(II). Further, $\lambda$ tends to zero, makes MD estimation of CG(II) little unstable and dependent on the choice of initial value as *optim* function often returns a local optima in R, as mentioned earlier.

In summary, K-S criterion suggests that CG(II) gives better fit to the raw data. However, if we consider group data with the class intervals as shown in Table 6.2., CG(I) gives a better fit as per the chi-square criterion. Of course, chi-square statistic depends on the choice of class intervals and can vary a bit. This example shows that there is no clear winner between the CG(I) and the geometric distribution, but under certain situations as mentioned earlier, CG(I) may give better fit to the data set than the geometric distribution.

Table 6.1. ML and MD estimates of the parameters of CG(I) and CG(II) and related statistics for the air-crash casualty data.

| Distribution | Method | Estimates $(\hat{\lambda}, \hat{q})$ | SEs $(\hat{\lambda}, \hat{q})$ | AIC | BIC | K-S | p-value |
|---|---|---|---|---|---|---|---|
| CG(I) | ML | (1.1168; 0.9606) | (2.4368; 0.0004) | 347.33 | 343.33 | 0.1319 | 0.3941 |
| | MD | (0.1709; 0.9429) | x | x | x | 0.1280 | 0.4310 |
| CG(II) | ML | (0.0100; 0.9469) | (0.9790; 0.0002) | 348.12 | 344.12 | 0.1050 | 0.6778 |
| | MD | (0.0099; 0.9410) | x | x | x | 0.1206 | 0.5063 |

x: Not relevant

Table 6.2. Observed and expected frequencies of the air-crash casualty data.

| No of fatalities | Observed frequency | expected frequency CG(I) MLE | CG(I) MDE | geometric MLE $(\hat{q} = 0.9469)$ | geometric MDE $(\hat{q} = 0.9396)$ | CG(II) MLE | CG(II) MDE |
|---|---|---|---|---|---|---|---|
| 0-9 | 23 | 20.23 | 20.48 | 18.48 | 20.40 | 18.46 | 19.99 |
| 10-19 | 8 | 9.90 | 10.73 | 10.72 | 10.94 | 10.72 | 10.92 |
| 20-29 | 5 | 5.38 | 5.76 | 6.22 | 5.87 | 6.22 | 5.95 |
| 30-39 | 3 | 3.13 | 3.14 | 3.60 | 3.15 | 3.61 | 3.25 |
| 40-49 | 2 | 1.91 | 1.73 | 2.09 | 1.69 | 2.09 | 1.77 |
| 50 or More | 3 | 3.45 | 2.14 | 2.89 | 1.95 | 2.88 | 2.12 |
| Total | 44 | | | | | | |
| Chi Square | | 0.8390 | 1.492 | 2.144 | 1.8769 | 2.1634 | 1.7859 |
| statistic | p-value | 0.8401 | 0.6839 | 0.7093 | 0.7584 | 0.5392 | 0.6180 |

# 7 Conclusion

In this paper, we have introduced two discrete distributions involving two parameters called the compounded geometric (CG) distributions and named as CG(I) and CG(II) respectively. Some statistical and reliability properties of both the distributions are derived with plots of the pmfs and the hazard rates. Estimation by maximum likelihood method and by distance minimization method are discussed. Finally, both the distributions are fitted with a real data set on casualties from plane crashes which show that the distributions have potential to be used in various applications.

Intuitively, as $\lambda$ approaches 0, a zero-truncated Poisson distribution becomes degenerate and assumes value 1 with probability 1. In such a case, subgroup size will also be a degenerate random variable which will assume value 1 with probability 1. When only one value is observed, eventually maximum and minimum is the same value and therefore both CG(I) and CG(II) will be same, as $\lambda$ approaches zero. In fact, in such a case, we will only observe a sequence of geometric random variables and the information on subgroup size will be redundant. That is, as $\lambda$ approaches 0, both CG(I) and CG(II) converge its parent geometric distribution. Further, when $\lambda$ is large enough and $p$ is also moderately large, most likely we shall have larger subgroup sizes with a few observations as 0. Obviously, the minimum will be 0 for most of the subgroups. As a consequence, CG(I) distribution will degenerate at 0. Although, this phenomenon is not observed in CG(II) for large $\lambda$.

The present work leaves two interesting future research problems. Note that CG(I) is DFR while CG(II) is IFR. It will be interesting to study the nature and the changing pattern of the compounded distribution obtained by any arbitrary order statistics in place of maximum or minimum. Another problem of worth considering is the simultaneous use of maximum and minimum to develop a bivariate geometric Poisson distribution.

**Acknowledgements:** Authors are grateful to Mr. Richard Kebabjian, owner of the website www.planecrashinfo.com for his kind permission to access the data.

## References

1. Adams, G. and Watson, R. (1989). A discrete time parametric model for the analysis of failure time data, *Australian Journal of Statistics*, **31**, pp. 365-384.

2. Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology*, **13**, S. Leinhardt, ed., Jossey-Bass, San Francisco, pp. 61-98.

3. Bartoszewicz, J. (2001). Stochastic comparisons of random minima and maxima from life distributions, *Statistics and Probability Letters*, **55**, pp. 107-112.

4. Belzunce, F., Ortega, E. M. and Ruiz, J. M. (2009). Aging properties of a discrete-time failure and repair model, *IEEE Transactions on Reliability*, **58(1)**, pp. 161-171.

5. Blyth, C. R. (1970). On the inference and decision models of statistics, *The Annals of Mathematical Statistics*, **41(3)**, pp. 1034-1058.

6. Boos, D. D. (1982). Minimum Anderson Darling Estimation, *Communications in Statistics-Theory and Methods*, **11(24)**, pp. 2747-2774.

7. Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization, *SIAM Journal of Scientific Computing*, **16**, pp. 1190-1208.

8. Drossos, C. A. and Philippou, A. N. (1980). A note on minimum distance estimates, *Annals of the Institute of Statistical Mathematics*, **32(1)**, pp. 121-123.

9. Fahrmeir, L. (1994). Dynamic modelling and penalized likelihood estimation for discrete time survival data, *Biometrika*, **81**, pp. 317-330.

10. Fahrmeir, L. and Knorr-Held, L. (1997). Dynamic discrete-time duration models: estimation via Markov Chain Monte Carlo, *Sociological Methodology*, **27(1)**, pp. 417-452.

11. Fahrmeir. L. and Wagenpfeil, S. (1996). Smoothing hazard functions and time-varying effects in discrete duration and competing risk models, *Journal of the American Statistical Association*, **91**, pp. 1584-1594.

12. Hamerle. A. (1986). Regression analysis for discrete event history or failure time data, *Statistical Papers*, **27**, pp. 207-225.

13. Hanselmann, A., Schrempf, O.C. and Hanebeck, U.D. (2007). Optimal parametric density estimation by minimizing an analytic distance measure, *Information Fusion, 10th International Conference on Digital Object Identifier*: 10.1109/ICIF.2007.4408100, pp. 1-8.

14. Jiang, S. and Kumar, R. (2001). Failure diagnosis of discrete event systems with linear-time temporal logic fault specifications, *IEEE Transactions on Automatic Control*, **49**, pp. 128-133.

15. Khalique, A. (1989). On discrete failure-time distributions, *Reliability Engineering and System Safety*, **25(2)**, pp. 99-107

16. Klar B. (1999). Model selection for a family of discrete failure time distribution, *Australian and New Zealand Journal of Statistics*, **41**, pp. 337-352.

17. Padgett, W.J. and Spurrier, J.D. (1985). On discrete failure models, *IEEE Transactions on Reliability*, **R-34**, pp. 253-256.

18. Parr W.C. and Schucany W. R. (1980). Minimum distance and robust estimation, *Journal of the American Statistical Association*, **75**, pp. 616-624.

19. Patil, P. N. and Bagkavos, D. (2012). Semiparametric smoothing of discrete failure time data, *Biometrical Journal; Special Issue: Survival and Event History Analysis*, **54(1)**, pp. 5-19.

20. R Development Core Team (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

21. Rocha-Martinez, J. M. and Shaked, M. (1995). A discrete-time model of failure and repair, *Applied Stochastic Models and Data Analysis*, **11(2)**, pp. 167-180.

22. Salvia, A.A and Bollinger, R.C. (1982). On discrete hazard functions, *IEEE Transaction on Reliability*, **R-31**, pp. 458-459.

23. Scheike, T. H. and Jensen, T. K. (1997). A discrete survival model with random effects: An application to time to pregnancy, *Biometrics*, **53**, pp. 349-360.

24. Shaked, M. and Shanthikumar, G. J. (2007). Stochastic Orders, Springer Series in Statistics.

25. Shaked, M. and Wong, T. (1997). Stochastic comparisons of random minima and maxima, *Journal of Applied Probability*, **34**, pp. 420-425.

26. Shis, J. H. (1998). Modeling multivariate discrete failure time data, *Biometrics*, **54**, pp. 1115-1128.

27. Singh, B., Sharma, K. K. and Kumar, A. (2009). Analyzing the dynamic system model with discrete failure time distribution, *Statistical Methods and Applications*, **18(4)**, pp. 521-542.

28. Weber, M. D., Leemis,L. M. and Kincaid R. K. (2006). Minimum Kolmogorov-Smirnov test statistic parameter estimates, *Journal of Statistical Computation and Simulation*, **76**, pp. 196-206.

29. Wolfowitz, J. (1957). The minimum distance method, *The Annals of Mathematical Statistics*, **28(1)**, pp. 75-88.

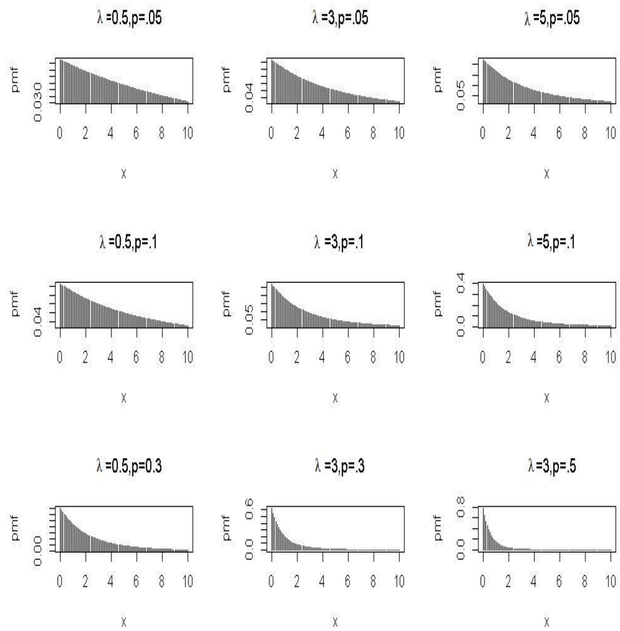


**Figure 2.A : pmf Plot for CG(I)**

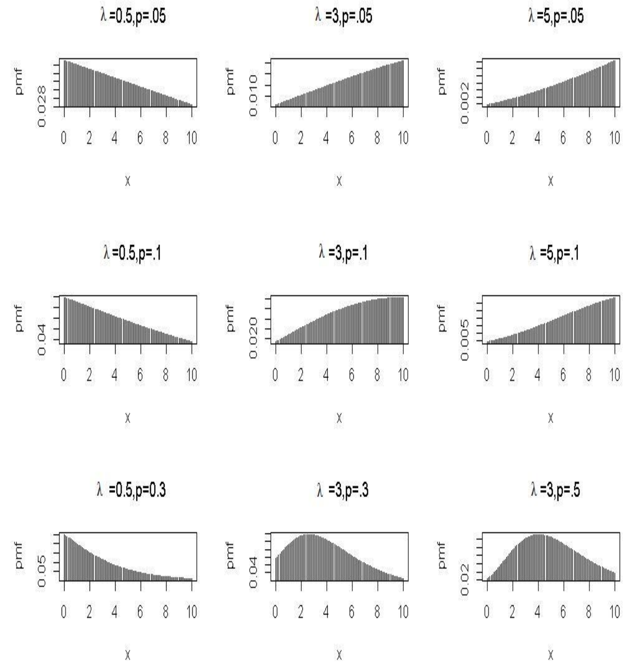


**Figure 2.B : pmf Plot for CG(II)**

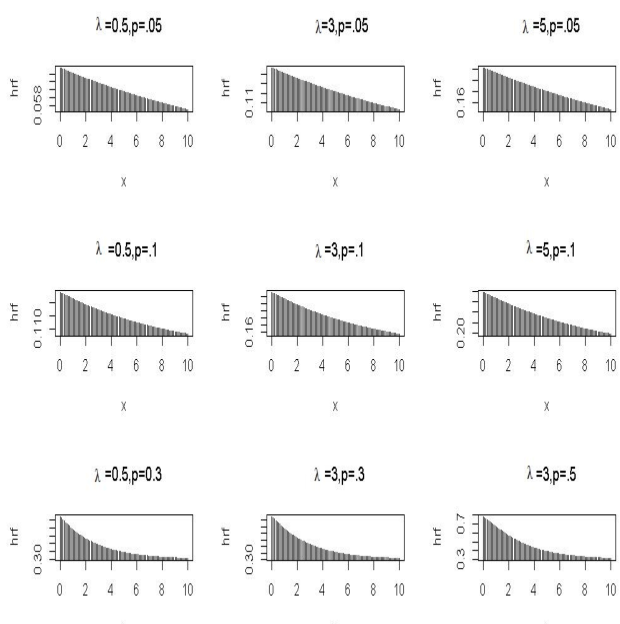


**Figure 3.3.A : hrf Plot for CG(I)**

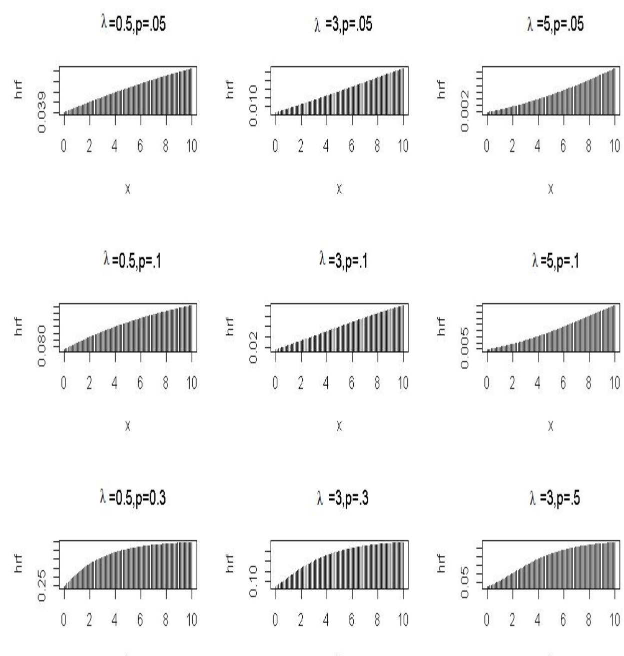


**Figure 3.3.B : hrf Plot for CG(II)**

TABLE 5.1. Comparison of MLEs and MDEs of parameters of CG(I) Distribution Based on Monte Carlo Simulation

| $\lambda$ | $q(=1-p)$ | n | Means of MLEs $(\hat{\lambda}, \hat{q})$ | SD of MLEs $(\hat{\lambda}, \hat{q})$ | Means of MDEs $(\hat{\lambda}, \hat{q})$ | SD of MDEs $(\hat{\lambda}, \hat{q})$ |
|---|---|---|---|---|---|---|
| 3 | 0.91 | 10 | 3.5152; 0.8879 | 3.3417; 0.0653 | 5.5364; 0.9020 | 4.1603; 0.0732 |
| | | 25 | 3.8628; 0.9053 | 3.0719; 0.0458 | 5.5674; 0.9176 | 4.0173; 0.0513 |
| | | 50 | 4.3465; 0.9144 | 3.2997; 0.0422 | 5.2432; 0.9210 | 3.6682; 0.0434 |
| | | 100 | 4.6841; 0.9208 | 3.2724; 0.0389 | 4.7167; 0.9205 | 3.0466; 0.0380 |
| | | 500 | 3.9545; 0.9202 | 1.8994; 0.0287 | 3.5669; 0.9135 | 1.5877; 0.0278 |
| | | 1000 | 3.6277; 0.9178 | 1.3602; 0.0238 | 3.2520; 0.9101 | 1.1383; 0.0233 |
| | | | | | | |
| | 0.93 | 10 | 4.5607; 0.9227 | 4.2889; 0.0529 | 5.8978; 0.9282 | 4.1893; 0.0564 |
| | | 25 | 4.1341; 0.9286 | 3.4895; 0.0362 | 5.5686; 0.9377 | 3.8405; 0.0381 |
| | | 50 | 4.3048; 0.9329 | 3.3057; 0.0324 | 5.0032; 0.9396 | 3.1909; 0.0310 |
| | | 100 | 4.4150; 0.9369 | 3.0579; 0.0299 | 4.5138; 0.9394 | 2.5124; 0.0265 |
| | | 500 | 3.6763; 0.9362 | 1.5861; 0.0201 | 3.6910; 0.9358 | 1.3866; 0.0198 |
| | | 1000 | 3.4132; 0.9343 | 1.1167; 0.0166 | 3.4241; 0.9328 | 1.1495; 0.0188 |
| | | | | | | |
| | 0.95 | 10 | 5.4152; 0.9493 | 4.8003; 0.0389 | 6.1410; 0.9506 | 4.0187; 0.0408 |
| | | 25 | 4.6865; 0.9487 | 4.3336; 0.0289 | 5.4530; 0.9560 | 3.4835; 0.0264 |
| | | 50 | 4.1451; 0.9476 | 3.6769; 0.0260 | 4.6839; 0.9554 | 2.7139; 0.0218 |
| | | 100 | 3.9534; 0.9487 | 3.1502; 0.0235 | 4.2185; 0.9546 | 2.2051; 0.0193 |
| | | 500 | 3.2983; 0.9496 | 1. 4487; 0.0157 | 3.5754; 0.9530 | 1.2943; 0.0143 |
| | | 1000 | 3.1852; 0.9499 | 1.0651; 0.0123 | 3.3914; 0.9523 | 0.9763; 0.0119 |
| | | | | | | |
| | 0.97 | 10 | 4.4751; 0.9648 | 4.3229; 0.0229 | 5.3507; 0.9683 | 3.1849; 0.0245 |
| | | 25 | 2.8779; 0.9592 | 2.8595; 0.0181 | 4.4408; 0.9707 | 2.4479; 0.0157 |
| | | 50 | 2.5999; 0.9600 | 1.9078; 0.0161 | 3.8418; 0.9701 | 1.8963; 0.0128 |
| | | 100 | 2.7644; 0.9631 | 1.6344; 0.0143 | 3.2177; 0.9683 | 1.4353; 0.0097 |
| | | 500 | 3.0533; 0.9681 | 1.1400; 0.0094 | 2.6935; 0.9664 | 0.5724; 0.0050 |
| | | 1000 | 3.1050; 0.9694 | 0.8719; 0.0069 | 2.6326; 0.9662 | 0.3643; 0.0036 |
| | | | | | | |
| 5 | 0.91 | 10 | 3.0567; 0.8551 | 2.5542; 0.0724 | 4.8463; 0.8289 | 3.8083; 0.1015 |
| | | 25 | 4.1243; 0.8866 | 3.0995; 0.0581 | 5.2761; 0.8620 | 3.8863; 0.0797 |
| | | 50 | 5.0045; 0.9040 | 3.3867; 0.0513 | 5.5391; 0.8743 | 3.9589; 0.0712 |
| | | 100 | 5.4823; 0.9160 | 3.2057; 0.0444 | 5.8234; 0.8861 | 3.8355; 0.0655 |
| | | 500 | 5.3046; 0.9253 | 1.9501; 0.0296 | 5.0579; 0.8886 | 2.6222; 0.0536 |
| | | 1000 | 5.1239; 0.9269 | 1.3249; 0.0220 | 4.7715; 0.8899 | 1.9996; 0.0472 |
| | | | | | | |
| | 0.93 | 10 | 3.0771; 0.8572 | 2.6046; 0.0762 | 5.3253; 0.8714 | 4.1498; 0.0871 |
| | | 25 | 4.1413; 0.8868 | 3.1551; 0.0594 | 5.7998; 0.8971 | 4.2036; 0.0656 |
| | | 50 | 4.9511; 0.9045 | 3.3372; 0.0511 | 6.0119; 0.9071 | 4.1283; 0.0570 |
| | | 100 | 5.5958; 0.9180 | 3.2618; 0.0445 | 6.0426; 0.9128 | 3.8809; 0.0520 |
| | | 500 | 5.2424; 0.9255 | 1.9181; 0.0292 | 5.1000; 0.9142 | 2.4684; 0.0415 |
| | | 1000 | 5.1081; 0.9274 | 1.3282; 0.0222 | 4.9863; 0.9176 | 1.9409; 0.0363 |
| | | | | | | |
| | 0.95 | 10 | 3.8505; 0.9014 | 3.6440; 0.0615 | 6.0502; 0.9178 | 4.3438; 0.0651 |
| | | 25 | 4.2405; 0.9176 | 3.3466; 0.0449 | 5.9956; 0.9303 | 4.1376; 0.0459 |
| | | 50 | 4.9074; 0.9292 | 3.4374; 0.0386 | 5.7054; 0.9331 | 3.8344; 0.0396 |
| | | 100 | 5.3738; 0.9372 | 3.3306; 0.0342 | 5.2293; 0.9336 | 3.2541; 0.0352 |
| | | 500 | 5.1517; 0.9445 | 1.9411; 0.0228 | 4.4308; 0.9339 | 1.9226; 0.0275 |
| | | 1000 | 5.1241; 0.9474 | 1.3914; 0.0167 | 4.2525; 0.9337 | 1.5576; 0.0253 |
| | | | | | | |
| | 0.97 | 10 | 5.8701; 0.9520 | 4.9368; 0.0394 | 6.4033; 0.9532 | 4.1089; 0.0397 |
| | | 25 | 5.1786; 0.9516 | 4.5893; 0.0303 | 5.8148; 0.9595 | 3.4868; 0.0263 |
| | | 50 | 5.0388; 0.9530 | 4.2304; 0.0276 | 5.2442; 0.9597 | 2.9666; 0.0225 |
| | | 100 | 4.7922; 0.9542 | 3.6623; 0.0259 | 4.8486; 0.9599 | 2.4883; 0.0204 |
| | | 500 | 4.4794; 0.9601 | 1.9740; 0.0186 | 4.4268; 0.9615 | 1.5971; 0.0158 |
| | | 1000 | 4.6837; 0.9642 | 1.5167; 0.0142 | 4.4504; 0.9635 | 1.2659; 0.0125 |

TABLE 5.2. Comparisons of MLEs and MDEs of parameters of CG(II) Distribution Based on Monte Carlo Simulation

| λ | $q(=1-p)$ | n | Means of MLEs $(\hat{\lambda},\hat{q})$ | SD of MLEs $(\hat{\lambda},\hat{q})$ | Means of MDEs $(\hat{\lambda},\hat{q})$ | SD of MDEs $(\hat{\lambda},\hat{q})$ |
|---|---|---|---|---|---|---|
| 3 | 0.91 | 10 | 4.1225; 0.8991 | 2.6322; 0.0298 | 3.8843; 0.8920 | 2.8639; 0.0358 |
| | | 25 | 3.4070; 0.9062 | 1.5063; 0.0186 | 3.4298; 0.9026 | 1.8148; 0.0213 |
| | | 50 | 3.1741; 0.9086 | 0.9799; 0.0130 | 3.1557; 0.9069 | 1.1485; 0.0150 |
| | | 100 | 3.0992; 0.9093 | 0.6587; 0.0090 | 3.0614; 0.9085 | 0.7668; 0.0105 |
| | | 500 | 3.0176; 0.9101 | 0.2922; 0.0040 | 2.9748; 0.9101 | 0.3224; 0.0046 |
| | | 1000 | 3.0093; 0.9100 | 0.2044; 0.0028 | 2.9742; 0.9102 | 0.2232; 0.0032 |
| | | | | | | |
| | 0.93 | 10 | 4.1668; 0.9203 | 2.6648; 0.0239 | 3.8593; 0.9150 | 2.8412; 0.0274 |
| | | 25 | 3.4209; 0.9262 | 1.4903; 0.0148 | 3.4364; 0.9236 | 1.7792; 0.0170 |
| | | 50 | 3.2257; 0.9276 | 0.9942; 0.0104 | 3.2188; 0.9264 | 1.1186; 0.0119 |
| | | 100 | 3.1266; 0.9285 | 0.6729; 0.0071 | 3.0845; 0.9282 | 0.7561; 0.0082 |
| | | 500 | 3.0446; 0.9294 | 0.2847; 0.0031 | 3.0022; 0.9293 | 0.3176; 0.0036 |
| | | 1000 | 3.0305; 0.9295 | 0.2049; 0.0022 | 2.9998; 0.9294 | 0.2198; 0.0024 |
| | | | | | | |
| | 0.95 | 10 | 4.1038; 0.9436 | 2.6594; 0.0175 | 3.9606; 0.9384 | 2.8863; 0.0208 |
| | | 25 | 3.4626; 0.9472 | 1.5207; 0.0107 | 3.4679; 0.9452 | 1.7489; 0.0123 |
| | | 50 | 3.1993; 0.9486 | 0.9612; 0.0073 | 3.2302; 0.9476 | 1.1330; 0.0085 |
| | | 100 | 3.1215; 0.9491 | 0.6585; 0.0052 | 3.0842; 0.9490 | 0.7458; 0.0057 |
| | | 500 | 3.0381; 0.9497 | 0.2892; 0.0023 | 3.0417; 0.9497 | 0.3229; 0.0025 |
| | | 1000 | 3.0252; 0.9498 | 0.2034; 0.0016 | 3.0404; 0.9498 | 0.2251; 0.0018 |
| | | | | | | |
| | 0.97 | 10 | 4.0492; 0.9659 | 2.6579; 0.0105 | 3.9121; 0.9626 | 2.7708; 0.0128 |
| | | 25 | 3.3954; 0.9682 | 1.5169; 0.0066 | 3.4606; 0.9670 | 1.8124; 0.0074 |
| | | 50 | 3.1348; 0.9691 | 0.9575; 0.0045 | 3.2329; 0.9685 | 1.1103; 0.0051 |
| | | 100 | 3.0486; 0.9695 | 0.6620; 0.0031 | 3.0995; 0.9694 | 0.7425; 0.0036 |
| | | 500 | 2.9631; 0.9699 | 0.2849; 0.0014 | 3.0431; 0.9698 | 0.3228; 0.0015 |
| | | 1000 | 2.9574; 0.9699 | 0.2040; 0.0010 | 3.0168; 0.9698 | 0.2274; 0.0010 |
| | | | | | | |
| 5 | 0.91 | 10 | 6.1029; 0.9033 | 3.0514; 0.0242 | 5.5673; 0.9010 | 3.1964; 0.0275 |
| | | 25 | 5.7322; 0.9055 | 2.1306; 0.0160 | 5.6174; 0.9050 | 2.4589; 0.0179 |
| | | 50 | 5.3470; 0.9076 | 1.3879; 0.0111 | 5.4395; 0.9064 | 1.7268; 0.0129 |
| | | 100 | 5.1908; 0.9083 | 0.8946; 0.0076 | 5.2374; 0.9083 | 1.1386; 0.0090 |
| | | 500 | 5.0476; 0.9093 | 0.3782; 0.0034 | 5.1006; 0.9094 | 0.4661; 0.0040 |
| | | 1000 | 5.0330; 0.9094 | 0.2676; 0.0024 | 5.0747; 0.9096 | 0.3230; 0.0028 |
| | | | | | | |
| | 0.93 | 10 | 6.1076; 0.9244 | 3.0669; 0.0193 | 5.5820; 0.9222 | 3.1538; 0.0217 |
| | | 25 | 5.6716; 0.9270 | 2.1158; 0.0126 | 5.6753; 0.9253 | 2.5163; 0.0147 |
| | | 50 | 5.3022; 0.9285 | 1.3653; 0.0088 | 5.4377; 0.9267 | 1.6739; 0.0100 |
| | | 100 | 5.1504; 0.9292 | 0.9116; 0.0062 | 5.2570; 0.9279 | 1.1361; 0.0072 |
| | | 500 | 5.0053; 0.9299 | 0.3755; 0.0027 | 5.0914; 0.9291 | 0.4653; 0.0030 |
| | | 1000 | 4.9965; 0.9299 | 0.2705; 0.0019 | 5.0733; 0.9292 | 0.3204; 0.0022 |
| | | | | | | |
| | 0.95 | 10 | 6.1885; 0.9459 | 3.0459; 0.0137 | 5.6378; 0.9438 | 3.1602; 0.0159 |
| | | 25 | 5.7006; 0.9481 | 2.1215; 0.0088 | 5.6799; 0.9463 | 2.4927; 0.0108 |
| | | 50 | 5.3788; 0.9489 | 1.3933; 0.0062 | 5.4157; 0.9479 | 1.6824; 0.0074 |
| | | 100 | 5.2127; 0.9494 | 0.9086; 0.0043 | 5.0553; 0.9496 | 0.4192; 0.0021 |
| | | 500 | 5.0671; 0.9500 | 0.3752; 0.0018 | 5.0553; 0.9496 | 0.4192; 0.0020 |
| | | 1000 | 5.0509; 0.9500 | 0.2662; 0.0013 | 5.0331; 0.9497 | 0.3153; 0.0015 |
| | | | | | | |
| | 0.97 | 10 | 6.2291; 0.9675 | 3.0642; 0.0084 | 5.4861; 0.9662 | 3.0329; 0.0096 |
| | | 25 | 5.7601; 0.9685 | 2.1467; 0.0054 | 5.4998; 0.9679 | 2.2705; 0.0063 |
| | | 50 | 5.4016; 0.9692 | 1.3888; 0.0038 | 5.4111; 0.9688 | 1.7154; 0.0044 |
| | | 100 | 5.2049; 0.9696 | 0.8984; 0.0026 | 5.3095; 0.9692 | 1.1223; 0.0031 |
| | | 500 | 5.0712; 0.9698 | 0.3787; 0.0011 | 5.1219; 0.9698 | 0.4827; 0.0013 |
| | | 1000 | 5.0458; 0.9699 | 0.2663; 0.0008 | 5.0811; 0.9699 | 0.3312; 0.0009 |

## Appendix

**Proof of Result 3.1:** From (2.6), $r$th order raw moment $(\mu_r^{'})$ of CG(I) can be obtained as

$$E(U^r) = \frac{1}{e^{\lambda}-1}\sum_{u=0}^{\infty} u^r (e^{\lambda q^u} - e^{\lambda q^{u+1}})$$

Considering partial sum of the infinite series $\sum_{u=0}^{\infty} u^r(e^{\lambda q^u} - e^{\lambda q^{u+1}})$, it can be easily shown that the series diverges. As a consequence, for CG(I), moments of order $r(\geq 1)$ and hence, MRL function, do not exist for finite $\lambda$ and for $q$ (or $p$) $\in (0,1)$. Similar conclusion can be drawn easily for CG(II).

**Proof of Result 3.2.1:** Using (3.3), first quartile $(Q_1)$, second quartile or median $(\tilde{\mu})$ and third quartile $(Q_3)$ of CG(I) can be derived as $Q_1 = \frac{1}{\ln(q)} \ln\left(\frac{1}{\lambda}\ln((3/4)e^{\lambda} + 1/4)\right) - 1$, $\tilde{\mu} = \ln\left(\frac{1}{\lambda}\ln((1/2)e^{\lambda} + 1/2)\right) - 1$ and $Q_3 = \frac{1}{\ln(q)} \ln\left(\frac{1}{\lambda}\ln((1/4)e^{\lambda} + 3/4)\right) - 1$, resulting in $IQR = Q_3 - Q_1 = \frac{1}{\ln(q)} \ln\left(\frac{\ln\left(\frac{1}{4}e^{\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{\lambda}+\frac{1}{4}\right)}\right)$ and $b_{1X} = \frac{(Q_3-Q_2)-(Q_2-Q_1)}{(Q_3-Q_1)} = \frac{\ln\left(\frac{\ln\left(\frac{1}{4}e^{\lambda}+\frac{3}{4}\right)\cdot\ln\left(\frac{3}{4}e^{\lambda}+\frac{1}{4}\right)}{\left(\ln\left(\frac{1}{2}e^{\lambda}+\frac{1}{2}\right)\right)^2}\right)}{\ln\left(\frac{\ln\left(\frac{1}{4}e^{\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{\lambda}+\frac{1}{4}\right)}\right)}$.

**Proof of Result 3.2.2:** Using (3.5), first quartile $(Q_1)$, second quartile or median $(\tilde{\mu})$ and third quartile $(Q_3)$ of CG(II) can be derived as $Q_1 = \frac{1}{\ln(q)} \ln\left(\frac{-1}{\lambda}\ln((3/4)e^{-\lambda} + 1/4)\right) - 1$, $\tilde{\mu} = \ln\left(\frac{-1}{\lambda}\ln((1/2)e^{-\lambda} + 1/2)\right) - 1$ and $Q_3 = \frac{1}{\ln(q)} \ln\left(\frac{-1}{\lambda}\ln((1/4)e^{-\lambda} + 3/4)\right) - 1$, resulting in $IQR = Q_3 - Q_1 = \frac{1}{\ln(q)} \ln\left(\frac{\ln\left(\frac{1}{4}e^{-\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{-\lambda}+\frac{1}{4}\right)}\right)$ and $b_{1X} = \frac{(Q_3-Q_2)-(Q_2-Q_1)}{(Q_3-Q_1)} = \frac{\ln\left(\frac{\ln\left(\frac{1}{4}e^{-\lambda}+\frac{3}{4}\right)\cdot\ln\left(\frac{3}{4}e^{-\lambda}+\frac{1}{4}\right)}{\left(\ln\left(\frac{1}{2}e^{-\lambda}+\frac{1}{2}\right)\right)^2}\right)}{\ln\left(\frac{\ln\left(\frac{1}{4}e^{-\lambda}+\frac{3}{4}\right)}{\ln\left(\frac{3}{4}e^{-\lambda}+\frac{1}{4}\right)}\right)}$.

**Proof of Result 3.3.1:** Write $a(x) = \frac{e^{\lambda q^{x+1}}-1}{e^{\lambda q^x}-1}$. Let us for the time being pretend that $a(x)$ is continuous in its domain and it is sufficiently differentiable. Then, differentiating $a(x)$ with respect to $x$, we get

$$a^{'}(x) \overset{sign}{=} (e^{\lambda q^{x+1}} - 1)e^{\lambda q^x} - (e^{\lambda q^x} - 1)qe^{\lambda q^{x+1}} = y^{q+1} - y - qy^{q+1} + qy^q,$$

where $y = e^{\lambda q^x}$ and $y \geq 1$. Now, writing $b(y) = y^{q+1} - y - qy^{q+1} + qy^q$, we see, $b^{'}(y) \overset{sign}{=} q^2y^{q-1}(1-y)-(1-y^q) = c(y)$, say. Further, $c^{'}(y) \overset{sign}{=} qy^{q-1}d(y)$ where $d(y) = q(q-1)/y - q^2 + 1$, which is increasing in $y \in [1, e^{\lambda}]$, giving that $d(y) > 0$ for all $y$. Thus, $a'(x) > 0$ for all $x$, which implies that $a(x)$ is increasing in $x$ giving that $h(x)$ is decreasing $x$. Hence, CG(I) distribution is DFR.

**Proof of Result 3.3.2:** Expression (3.8) may be alternately written as $h^*(x; q, \lambda) = \frac{e^{\lambda p q^x}-1}{e^{\lambda q^x}-1}$.

As before, if we pretend same about the nature of the function $h^*$, then differentiating $h^*$ with respect to $x$, we get, $\frac{d}{dx}(h^*(x;q,\lambda)) = \lambda q^x \log q \left[e^{\lambda q^x(1+p)}(p-1) - e^{\lambda q^x}(1-pe^{-\lambda p q^x})\right] > 0$ as $\log q < 0$, $(p-1) < 0$ and $(1-pe^{-\lambda p q^x}) > 0$, $h^*(x;q,\lambda)$ is increasing in $x$. Hence, CG(II) distribution is IFR.