Mathematics and Statistics 10(6): 1218-1228, 2022
DOI: 10.13189/ms.2022.100608 http://www.hrpub.org

# Binomial-Geometric Mixture and Its Applications

**Hussein Eledum**[1,*], **Alaa R. El-Alosey**[2]

[1]Department of Statistics, Faculty of Science, University of Tabuk, KSA

[2]Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

*Received July 31, 2022; Revised September 27, 2022; Accepted October 11, 2022*

***Cite This Paper in the following Citation Styles***

***(a)**: [1] Hussein Eledum, Alaa R. El-Alosey, "Binomial-Geometric Mixture and Its Applications," Mathematics and Statistics, Vol.10, No.6, pp. 1218-1228, 2022. DOI: 10.13189/ms.2022.100608*

***(b)**: Hussein Eledum, Alaa R. El-Alosey (2022). Binomial-Geometric Mixture and Its Applications. Mathematics and Statistics, 10(6), 1218-1228. DOI: 10.13189/ms.2022.100608*



**Abstract** A mixture distribution is a combination of two or more probability distributions; it can be obtained from different distribution families or the same distribution families with different parameters. The underlying distributions may be discrete or continuous, so the resulting mixture probability distribution function should be a mass or density function. In the last few years, there has been great interest in the problem of developing a mixture distribution based on the binomial distribution. This paper uses the probability generating function method to develop a new two-parameter discrete distribution called a binomial-geometric (BG) distribution, a mixture of binomial distribution with the number of trials (parameter $n$) taken after a geometric distribution. The quantile function, moments, moment generating function, Shannon entropy, order statistics, stress-strength reliability and simulating the random sample are some of the statistical highlights of the BG distribution that are explored. The model's parameters are estimated using the maximum likelihood method. To examine the performance of the accuracy of point estimates for BG distribution parameters, the Monte Carlo simulation is performed with different scenarios. Finally, the BG distribution is fitted to two real lifetime count data sets from the medical field. As a result, the proposed BG distribution is an overdispersed right-skewed and can accommodate a constant hazard rate function. The proposed BG distribution is appropriate for modelling the overdispersed right-skewed real-life count data sets and it can be an alternative to the negative binomial and geometric distributions.

**Keywords** Mixture of Distributions, Binomial Distribution, Geometric Distribution, Hazard Function, Maximum Likelihood Estimation, Quantile Functions, Shannon Entropy, Stress-strength Parameter

## 1 Introduction

A mixture distribution is a statistical term that refers to a collection of two or more probability distributions. It can be used to model a statistical population with subpopulations, with the weights representing the percentages of each subpopulation in the total population and the subpopulation densities serving as the components of the mixture probability densities. The subpopulations can be univariate or multivariate and either be discrete or continuous probability distributions. In addition, the mixture distribution can be obtained from different distribution families or the same distribution families with different parameters. A mixture distribution may be appropriate for some data sets because distinct subsets of the total data set have different qualities that are better modeled individually. Moreover, the individual mixture components can be handled more effectively statistically than the entire mixture density, which makes the mixture distributions more tractable. Fisheries, agriculture, botany, economics, medicine, genetics, psychology, paleontology, electrophoresis, finance, communication theory, geology, and zoology are some fields where mixed distributions are applied.

In the last few years, there has been great interest in the problem of developing a mixture distribution based on the binomial distribution. Blischke [1] constructed a moment estimator for the parameters of the mixture of binomial distributions and obtained the covariance matrix of the joint asymptotic normal distribution of the estimators. Beta and Poisson mixture of binomial distribution was studied respectively by Schmittlein [2] and Roy et al [3]. Wood [4] introduced the mixture of binomial distribution using a cumulative distribution function G on [0,1]. Wood [5] investigated the geometric estimate of the binomial

mixture distribution. Roy et al. [6] described the binomial mixtures of some common standard distributions, including the Poisson, normal, log-normal, chi-square, F, t, beta, gamma, exponential, rectangular, and Erlang distributions, along with their various properties, they used a fact that the resulting distribution is the weighted average of the probability distribution with weights equal to the various terms of the binomial distribution.

Recently Zhu et al. [7] utilized a beta-binomial-Poisson mixture distribution to model the number of successes and number of binary trials at the same time, with an application to foraging success of adult and immature green-backed herons. Faddy and Smith [8] proposed an alternative scenario to [7] procedure by decomposing these components by modeling two processes, one for the number of attempts at foraging and another for success at each attempt. Shkedy et al. [9] introduced the hierarchical Binomial-Poisson assuming the number of responses to be a Poisson random variable, for the analysis of a crossover design for correlated binary data when the number of trials is dose-dependent. Grilli et al. [10] used a binomial finite mixture model to simulate the number of credits earned by undergraduates during their first year at the University of Florence's School of Economics. Knape et al. [11] assessed the sensitivity of binomial N-mixture models to overdispersion in abundance and detection using simulations and a case study. El-Alosey [12] derived the probability mass function of discrete mixtures of distributions by using the probability generating function method and used this idea to introduced binomial- exponential mixture. Abed Al-Kadim and AL-Hussani [13] introduced the binomial mixture of Erlang distribution based on transformation $p = e^{-\lambda}$ using moment method and Laplace transform. Very recently, the triple Binomial was introduced by Adnan and Kiser [14] using the triple-mixture distributions which are defined as multiplicative mixture of the three same distributions. Adnan et al. [15] introduced a new class of mixture distributions call power function mixtures of distributions.

In this paper, we introduce a new mixture of binomial distribution by mixing the binomial and the geometric distribution, which is called the binomial-geometric (BG) mixture. We drive the pmf of BG distribution by using the probability generating function of mixtures. We investigate some statistical properties of the proposed distribution and estimate its parameters using the maximum likelihood method. We conduct a Monte Carlo simulation to examine the performance of the accuracy of point estimates for BG distribution. Moreover, we fit the proposed BG distribution to two real lifetime count data sets from the medical field.

The remainder of this paper is structured as follows: The new proposed distribution BG is presented in section 2, some statistical properties such as quantile function, moment generating function, Shannon entropy, order statistics, stress-strength parameter and simulating of the random sample are demonstrated in section 3. Section 4 provides the maximum likelihood method to estimate BG mixture parameters. The BG distribution is applied to two real data sets in section 5. Finally, section 6 provides some concluding remarks.

### 1.1 Binomial and Geometric distributions

Consider a binomial random variable $X$ with parameters $n$ and $p$ with probability mass function (pmf) as:

$$\begin{aligned} &f_X(x;n,p) = \binom{n}{x}\ p^x\ \ q^{n-x} \\ &x = 0,1,2,\ldots,n \quad ,\ 0 \le p \le 1, \quad p+q=1 \end{aligned} \tag{1}$$

where $n \in \{0,1,2,\ldots\}$ is a nonnegative integer representing number of trials. The probability generating function (pgf) of the binomial random variable $X$ is

$$P_X(z;n,p) = E\left(z^x\right) = \ (q+pz)^n$$

The geometric distribution is a discrete probability distribution used to model the probability of experiencing a certain number of failures before encountering the first success in a sequence of Bernoulli trials. The pmf of the geometric distribution is given by

$$f_N(n;\theta) = \theta\left(1-\theta\right)^n\ ;\ n \in \{0,1,2,\ldots\}\ ,\ \ 0 \le \theta\ \le 1 \tag{2}$$

where $n$ represents the number of failures before the first success.
The mean and variance of the geometric distribution are given, respectively as

$$\begin{aligned} E(n) &= 1/\theta \\ Var(n) &= (1-\theta)/\theta^2 \end{aligned}$$

Note that for $\theta < 0.5$ the variance of the geometric distribution is overdispersed (the variance is greater than mean).

### 1.2 Binomial mixtures using Method of Generating Function

Suppose that the parameter $n$ of the binomial distribution in Eq.(1) is a random variable having a distribution with pmf $f_N(n,\theta)$, then the binomial mixture distribution can be obtained by using the probability generating function method as

$$f\left(z;\ p,\theta\right) = \sum_{n=0}^{\infty} P_X(z;n,p) f_{\,N}(n,\theta) \tag{3}$$

where $P_X(z;n,p)$ is pgf for the binomial distribution, while $p$, and $\theta$ are the parameters of the mixture distribution.

## 2 Binomial-Geometric distribution (BG)

The mathematical formulas as well as the graphs of the probability mass and cumulative distribution functions of the proposed binomial-geometric mixture are evaluated and discussed in this section. Moreover, this section also demonstrates the survival and hazard rate functions for the BG distribution.

## 2.1 Probability and cumulative distribution functions for the BG distribution

**Definition 2.1.** A random variable $X$ is said to have BG distribution if its probability mass function is given by

$$f\left(x; p, \theta\right) = \frac{\theta\, p^{x}\left(1-\theta\right)^{x}}{\left[1-q\left(1-\theta\right)\right]^{x+1}} \tag{4}$$

where, $x = 0, 1, 2, \ldots,\ 0 \leq p \leq 1,\ 0 \leq \theta \leq 1, p+q = 1$

**Definition 2.2.** A random variable $X$ is said to have BG distribution if its cumulative distribution function is defined by

$$F_X\left(x;\ p,\ \theta\right) = 1 - \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{x+1} \tag{5}$$

Note that $f\left(x; p, \theta\right)$ of the BG distribution in Eq.(4) is obtained by using Eq.(3) assuming that the parameter $n$ in a binomial distribution follows a geometric distribution in Eq.(2), that is

$$\begin{aligned} P\left(z;\ p,\ \theta\right) &= \sum_{n=0}^{\infty} P_X(z; n, p) f_N(n;\ \theta) \\ &= \theta \sum_{n=0}^{\infty} \left[q\left(1-\theta\right) + p\left(1-\theta\right) z\right]^{n} \\ &= \frac{\theta}{1-q(1-\theta)-p(1-\theta)z} \\ &= \frac{\theta}{\left(1-q\left(1-\theta\right)\right)\left(1-\frac{p(1-\theta)z}{1-q(1-\theta)}\right)} \\ &= \frac{\theta}{1-q(1-\theta)} \sum_{i=0}^{\infty} \left(\frac{p(1-\theta)z}{1-q(1-\theta)}\right)^{i} \end{aligned}$$

Thus the pmf of BG is the coefficient of $z^x$ in the pgf and is written as in Eq.(4).

**Lemma 2.1.** The limit of BG distribution function as $x \to -\infty$ is 0 and $x \to \infty$ is 1.

**Proof.**

$$\begin{aligned} F(\infty) &= \lim_{x\to\infty} \left\{1 - \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{x+1}\right\} \\ &= 1 - \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{\infty} \\ &= 1 \end{aligned}$$

because

$$\begin{aligned} 1-q(1-\theta) &= \theta + (1-\theta) - q(1-\theta) \\ &= \theta + (1-\theta)(1-q) \\ &= \theta + p(1-\theta) \end{aligned}$$

since $\theta > 0$, therefore, $\theta + p(1-\theta) > p(1-\theta)$ and $\frac{p(1-\theta)}{1-q(1-\theta)} < 1$. It follows that

$$\left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{\infty} = 0$$

and

$$\begin{aligned} F(-\infty) &= \lim_{x\to\infty} \left\{1 - \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{x+1}\right\} \\ &= 1 - \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{-\infty} \\ &= 1 - \left[\frac{1-q\left(1-\theta\right)}{p\left(1-\theta\right)}\right]^{\infty} \\ &= 0 \end{aligned}$$

because the quantity $\left[\frac{1-q(1-\theta)}{p(1-\theta)}\right]^{\infty}$ doesn't exceed 1. For example $F(-1) = 0$.

**Lemma 2.2.** $f(x)$ of Eq.(4) is a probability mass function

**Proof.**

To prove $f(x)$ is a pmf, we need to prove $f(x) \geq 0$ and $\sum_{x=0}^{\infty} f(x) = 1$

We know that the limit of BG probability mass function as $x \to 0$ is $\frac{\theta}{1-q(1-\theta)}$, since $0 \leq q \leq 1,\ 0 \leq \theta \leq 1$ then $\frac{\theta}{1-q(1-\theta)} \geq 0$. It follows that $f(x) \geq 0$.

$$\begin{aligned} \sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} \frac{\theta\, p^{x}\left(1-\theta\right)^{x}}{\left[1-q\left(1-\theta\right)\right]^{x+1}} \\ &= \sum_{x=0}^{\infty} \frac{\theta}{1-q(1-\theta)} \left[\frac{p(1-\theta)}{1-q(1-\theta)}\right]^{x} \\ &= \frac{\theta}{1-q(1-\theta)} \left[1 + \frac{p(1-\theta)}{1-q(1-\theta)} + \left[\frac{p(1-\theta)}{1-q(1-\theta)}\right]^{2} + \ldots\right. \end{aligned}$$

since $\frac{p(1-\theta)}{1-q(1-\theta)} < 1$, then

$$\begin{aligned} \sum_{x=0}^{\infty} f(x) &= \frac{\theta}{1-q(1-\theta)} \left[\frac{1}{1-\frac{p(1-\theta)}{1-q(1-\theta)}}\right] \\ &= \frac{\theta}{1-q(1-\theta)-p(1-\theta)} \\ &= \frac{\theta}{1-(p+q)(1-\theta)} \\ &= \frac{\theta}{\theta} \\ &= 1 \end{aligned}$$

Figures 1 and 2 demonstrate respectively the pmf and cdf of BG distribution for combination values of $p$ and $\theta$.

Looking at Figure 1, it is observed that the proposed distribution is a right-skewed and the pmf is a decreasing function. Moreover, number of zeros frequencies increase with the parameter $\theta$, whereas, when the parameter $p$ increases, the distribution expands to include a large number of $x$ values.

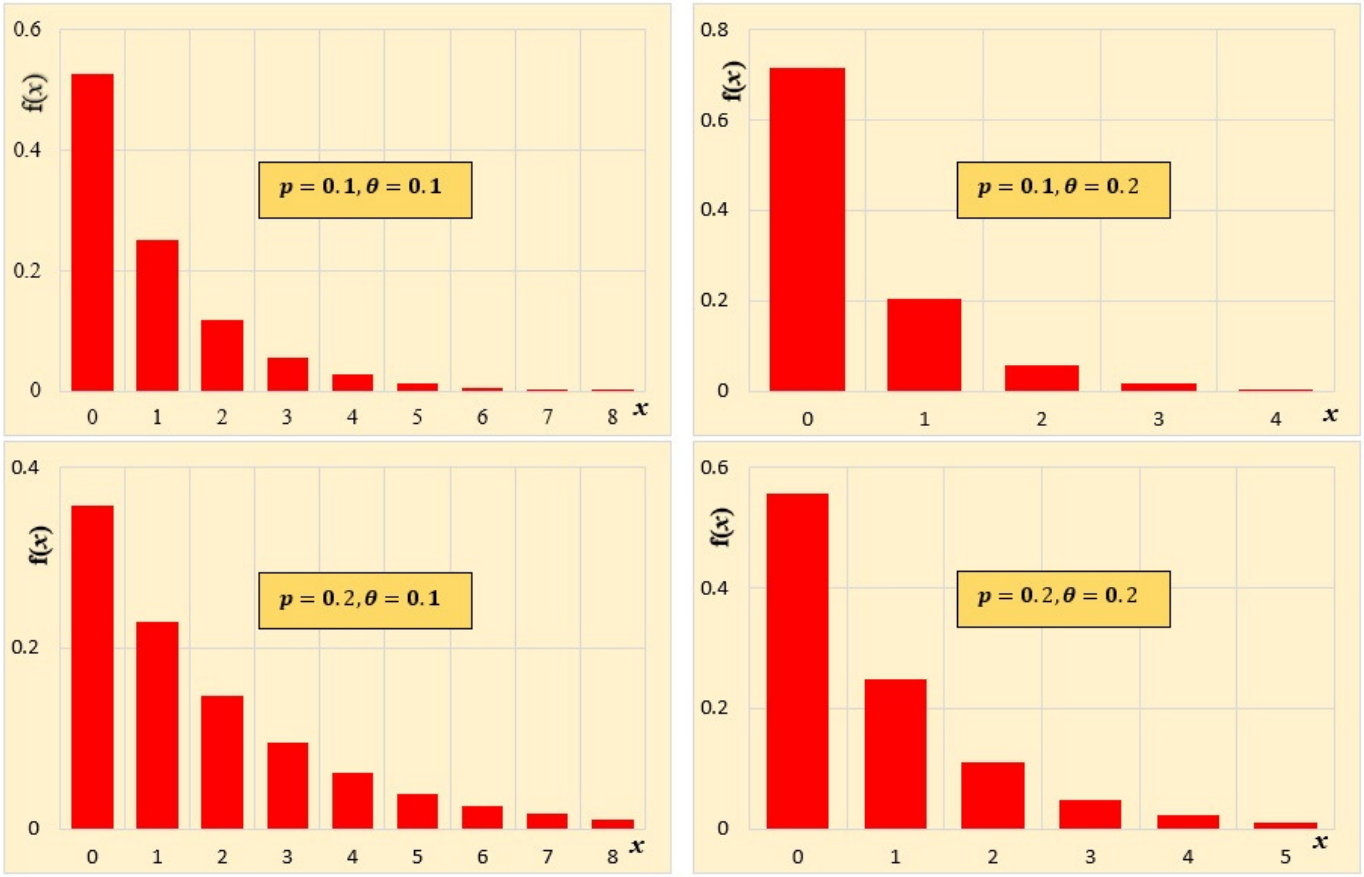


**Figure 1.** The pmfs of BG distribution for different values of $p$ and $\theta$

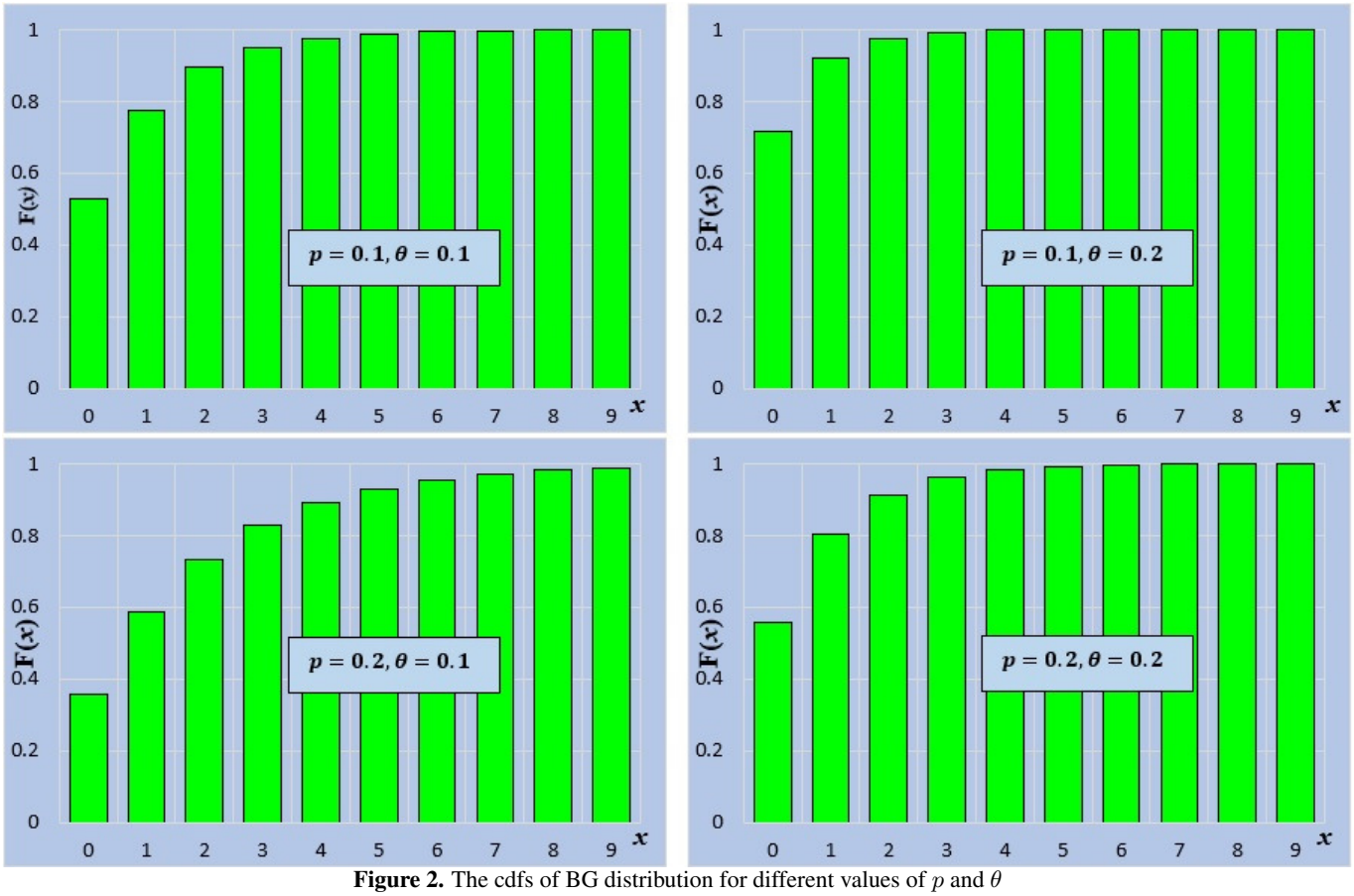


**Figure 2.** The cdfs of BG distribution for different values of $p$ and $\theta$

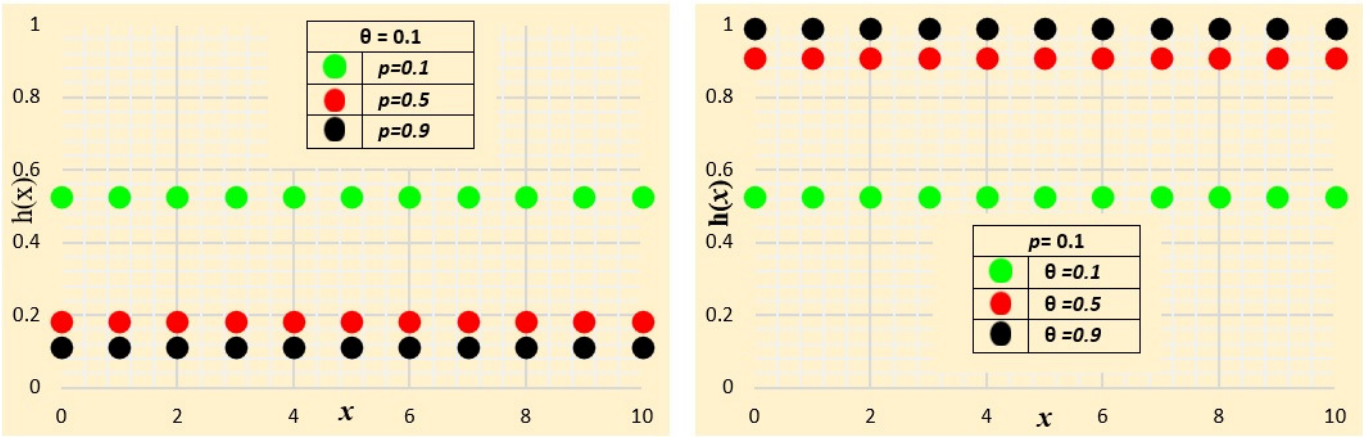


**Figure 3.** The hazard function of BG distribution for different values of $p$ and $\theta$

### 2.2 Survival and Hazard function for the BG distribution

**Definition 2.3.** The survival function of a random variable X having a binomial-geometric distribution is defined by

$$s_X\left(x;\ p,\ \theta\ \right) = \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^x \tag{6}$$

The survival function of Eq.(6) is obtained as

$$\begin{aligned} s_X(x;\ p,\theta) =& P\left(X \geq x\right) = P\left(X = x\right) + P\left(X > x\right) \\ =& P\left(X = x\right) + 1 - F_X\left(x;\ p,\ \theta\right) \\ =& 1 - F_X\left(x-1;\ p,\ \theta\right) \\ =& \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^x \end{aligned}$$

**Definition 2.4.** The hazard function of a random variable X having a binomial-geometric distribution is given as

$$h_X\left(x;\ p,\ \beta\ \right) = \frac{\theta}{1-q(1-\theta)}$$

Figure 3 as well as Table 1 demonstrate the behavior of the hazard rate function of BG for various values of $p$ and $\theta$ . From Figure 3 and Table 1, it is obvious that the hazard rate function of the proposed BG distribution is a constant. Moreover, it is growing up with $\theta$ and shrinking down with $p$.

**Table 1.** The hazard function of BG distribution for different values of $p$ and $\theta$

| | $\theta$=0.1 | $\theta$=0.25 | $\theta$=0.5 | $\theta$=0.75 | $\theta$=0.9 |
|---|---|---|---|---|---|
| $p$=0.1 | 0.5263 | 0.7692 | 0.9091 | 0.9677 | 0.989 |
| $p$=0.25 | 0.3077 | 0.5714 | 0.8 | 0.9231 | 0.973 |
| $p$=0.5 | 0.1818 | 0.4 | 0.6667 | 0.8571 | 0.9474 |
| $p$=0.75 | 0.129 | 0.3077 | 0.5714 | 0.8 | 0.9231 |
| $p$=0.9 | 0.1099 | 0.2703 | 0.5263 | 0.7692 | 0.9091 |

## 3 Distributional properties

In this section, we derive some statistical properties of BG distribution including quantile function, moment generating function, and some dispersion measures. Moreover, we also obtain some other statistical techniques involving Shannon entropy, order statistics, stress-strength parameter, and simulating the random sample.

### 3.1 Quantile and moment generating functions

**Theorem 3.1.** If $X$ is a binomial-geometric distribution with the parameters $p$ and $\theta$ then the $r^{th}$ quantile is given as

$$Q\left(r;\ p,\theta\ \right) = \frac{\log_2\left(\ 1-r\right)}{\log_2\left[p(1-\theta)\right] - \log_2\left[1-q(1-\theta)\right]} - 1 \tag{7}$$

and the median is

$$Q_{\ 0.5} = Q\left(r;p,\theta\ \right) = \frac{-1}{\log_2\left[p(1-\theta)\right] - \log_2\left[1-q(1-\theta)\right]} - 1$$

**Proof.** The $r^{th}$ quantile is obtained by inverting the cdf in Eq.(5) as follows

$$F_X\left(Q;p,\ \theta\right) = 1 - \left[\frac{p(1-\theta)}{1-q(1-\theta)}\right]^{Q+1}$$

Then $F_X^{-1}\left(r\right) = \min\{x \in R : F_X(x) \geq r\}$ that is,

$$\begin{aligned} 1 - \left(\frac{p(1-\theta)}{1-q(1-\theta)}\right)^{Q+1} &= r \\ \left(\frac{p(1-\theta)}{1-q(1-\theta)}\right)^{Q+1} &= r+1 \\ (Q+1)\log_2\left(\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right) &= \log_2\left(\ 1-r\right) \\ Q &= \frac{\log_2\left(\ 1-r\right)}{\log_2\left[p(1-\theta)\right] - \log_2\left[1-q(1-\theta)\right]} - 1 \end{aligned}$$

Thus, The $r^{th}$ quantile is

$$Q\left(r;\ p,\theta\ \right) = \frac{\log_2\left(\ 1-r\right)}{\log_2\left[p(1-\theta)\right] - \log_2\left[1-q(1-\theta)\right]} - 1$$

The median of BG distribution is computed by substituting by $r = \frac{1}{2}$ in Eq.(7).

**Theorem 3.2.** If $X$ is a binomial-geometric distribution with

the parameters $p$ and $\theta$ then the moment generating function is given as

$$M_X(t) = \frac{\theta}{1-q(1-\theta)-p(1-\theta)e^t} \tag{8}$$

and
The first moment about the origin (mean $\mu$) is

$$\mu_1 = \frac{p(1-\theta)}{\theta} \tag{9}$$

The second moment is

$$\mu_2 = \frac{p(1-\theta)}{\theta} + \frac{2\left[p(1-\theta)\right]^2}{\theta^2}$$

The variance is

$$\sigma^2 = \frac{p(1-\theta)}{\theta} + \frac{\left[p(1-\theta)\right]^2}{\theta^2} \tag{10}$$

The third moment about the origin is

$$\mu_3 = \frac{p(1-\theta)}{\theta} + \frac{6\left[p(1-\theta)\right]^2}{\theta^2} + \frac{6\left[p(1-\theta)\right]^3}{\theta^3}$$

The fourth moment about the origin is

$$\mu_4 = \frac{p(1-\theta)}{\theta} + \frac{14\left[p(1-\theta)\right]^2}{\theta^2} + \frac{36\left[p(1-\theta)\right]^3}{\theta^3} + \frac{24\left[p(1-\theta)\right]^4}{\theta^4}$$

The coefficient of variation is

$$C.V = \sqrt{\frac{1-q(1-\theta)}{p(1-\theta)}}$$

The Skewness is

$$\sqrt{\beta_1} = \frac{\left[\frac{p(1-\theta)}{\theta} + \frac{3[p(1-\theta)]^2}{\theta^2} + \frac{2[p(1-\theta)]^3}{\theta^3}\right]}{\left[\frac{p(1-\theta)}{\theta} + \frac{[p(1-\theta)]^2}{\theta^2}\right]^{\frac{3}{2}}} \tag{11}$$

The Kurtosis is

$$\beta_2 = \frac{\left[1 + \frac{10\,p(1-\theta)}{\theta} + \frac{18[p(1-\theta)]^2}{\theta^2} + \frac{4[p(1-\theta)]^3}{\theta^3}\right]}{\left[\frac{p(1-\theta)}{\theta} + \frac{2[p(1-\theta)]^2}{\theta^2} + \frac{[p(1-\theta)]^3}{\theta^3}\right]}$$

The index of dispersion is

$$\gamma = \frac{p+q\theta}{\theta}$$

**Proof.** The proof is straightforward.
**Corollary 3.1.** The BG distribution is the overdispersed, that is, the variance is always greater than the mean.
**Proof.** It is obvious from Eqs.(9) and (10) that

$$\sigma^2 = \mu_1 + \mu_1^2 > \mu$$

**Corollary 3.2.** The BG distribution is a right skewed.
**Proof.** Since $p$ and $\theta$ are positive, and $0 < \theta < 1$ then $\sqrt{\beta_1}$ of Eq.(11) is always greater than zero. Table 2 shows the mean and variance of the BG distribution for various combinations of $p$ and $\theta$. It is self-evident that, as $p$ increases, so do the mean and variance. In contrast as $\theta$ increases, the mean and variance decrease. As a result, most overdispersed right skews count data sets may be fitted using the proposed distribution's two parameters. Figure 4 displays the coefficient of variation, coefficient of Skewness, coefficient of Kurtosis, and the index of dispersion of the BG distribution for different values of $p$ and $\theta$. From Figure 4, it is abundantly clear that as $\theta$ increases, so do the coefficient of variation, skewness, and kurtosis, while these measures decrease as $p$ increases. On the other hand, the index of dispersion grows up with $p$ and drops down as $\theta$ increases.

### 3.2 Shannon entropy

A probabilistic measure of uncertainty or ignorance regarding the result of a random experiment, as well as a measure of reduction in that uncertainty, is known as statistical entropy. The Shannon entropy, which has been developed and utilized in a range of disciplines and contexts, is one of many entropy and information indices and defined as

$$H(X) = E\left(-\log\left[f(x)\right]\right)$$

The Shannon entropy for a random variable $X$ with a pmf of the BG distribution in Eq.(4) is computed as

$$\begin{aligned} H(X) &= -\sum_{x=0}^{\infty} f(x)\log[f(x)] \\ &= \frac{p(1-\theta)}{\theta}\log\left[\frac{1-q(1-\theta)}{p(1-\theta)}\right] - \log\left[\frac{\theta}{1-q(1-\theta)}\right] \end{aligned}$$

Table 3 as well as Figure 5 demonstrate Shannon entropy of the BG distribution for different values of $p$ and $\theta$. It is observed that the Shannon entropy increases with $p$ and decreases as $\theta$ increases.

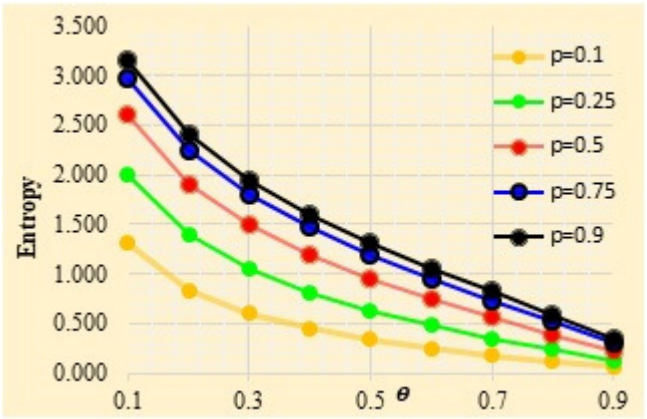


**Figure 5.** Shannon entropy of BG distribution for different values of $p$ and $\theta$

### 3.3 Order statistics for BG distribution

The most crucial fundamental tools for non-parametric statistics and inference are order statistics. They take a range of approaches to estimate and hypothesis testing problems. The

**Table 2.** The mean and variance of the BG distribution for diverse values of $p$ and $\theta$

| | $\theta$=0.1 | | $\theta$=0.25 | | $\theta$=0.5 | | $\theta$=0.75 | | $\theta$=0.9 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Variance | Mean | Variance | Mean | Variance | Mean | Variance | Mean | Variance |
| $p$=0.1 | 0.9 | 1.71 | 0.3 | 0.39 | 0.1 | 0.11 | 0.0333 | 0.0344 | 0.0111 | 0.0112 |
| $p$=0.25 | 2.25 | 7.3125 | 0.75 | 1.3125 | 0.25 | 0.3125 | 0.0833 | 0.0903 | 0.0278 | 0.0285 |
| $p$=0.5 | 4.5 | 24.75 | 1.5 | 3.75 | 0.5 | 0.75 | 0.1667 | 0.1944 | 0.0556 | 0.0586 |
| $p$=0.75 | 6.75 | 52.3125 | 2.25 | 7.3125 | 0.75 | 1.3125 | 0.25 | 0.3125 | 0.0833 | 0.0903 |
| $p$=0.9 | 8.1 | 73.71 | 2.7 | 9.99 | 0.9 | 1.71 | 0.3 | 0.39 | 0.1 | 0.11 |

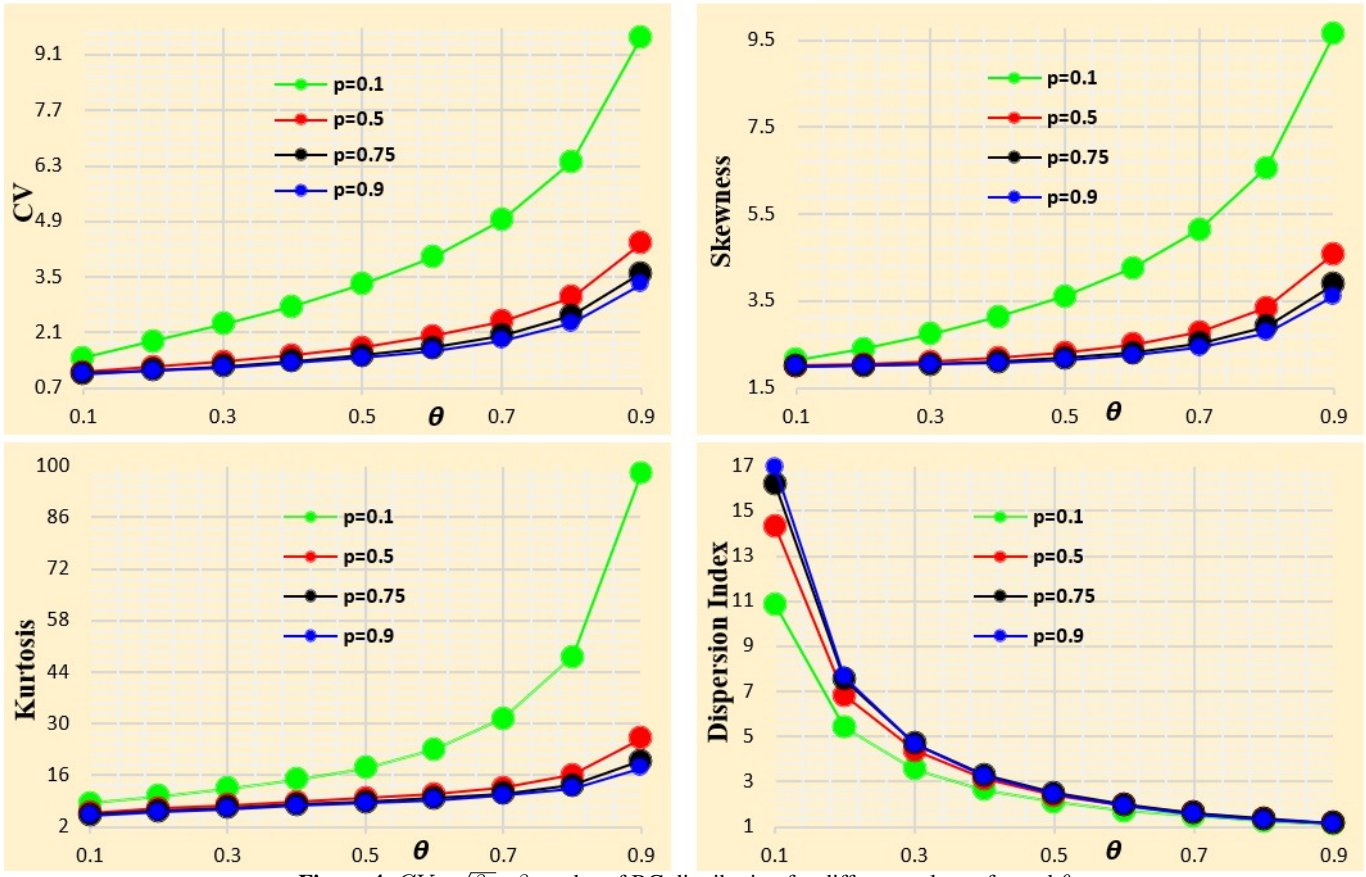


**Figure 4.** $CV.$, $\sqrt{\beta_1}$ , $\beta_2$ and $\gamma$ of BG distribution for different values of $p$ and $\theta$

**Table 3.** Shannon entropy of BG distribution for different values of $p$ and $\theta$

| $p$ \ $\theta$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 1.314 | 0.838 | 0.598 | 0.445 | 0.335 | 0.249 | 0.179 | 0.118 | 0.061 |
| 0.25 | 2.006 | 1.386 | 1.042 | 0.806 | 0.626 | 0.478 | 0.352 | 0.238 | 0.128 |
| 0.5 | 2.608 | 1.91 | 1.495 | 1.195 | 0.955 | 0.75 | 0.566 | 0.392 | 0.218 |
| 0.75 | 2.98 | 2.249 | 1.803 | 1.469 | 1.195 | 0.955 | 0.733 | 0.518 | 0.294 |
| 0.9 | 3.151 | 2.408 | 1.949 | 1.603 | 1.314 | 1.059 | 0.819 | 0.584 | 0.335 |

goal of this part is to establish some general BG distribution equations. Specifically, let $f_k(x; p,\ \theta)$ and $F_k(x; p,\ \theta)$ be the pmf and cdf of the $k^{th}$ order statistic of a random sample; $X_1, X_2, \ldots, X_n$; of size $n$, respectively, derived from BG distribution. In addition, we calculate the pmf of the maximum, minimum, and median order statistics.
The $k^{th}$ order statistic's pmf is

$$\begin{aligned} f_k\left(x; p\ , \theta\right) &= \frac{n!}{(k-1)!\,(n-k)!}\left[F\left(x; p,\ \theta\right)\right]^{k-1} \\ &\times \left[1 - F\left(x; p,\ \theta\right)\right]^{n-k} f\left(x; p,\ \theta\right) \\ &= \frac{n!}{(k-1)!\,(n-k)!}\sum_{j=0}^{k-1}(-1)^j \binom{k-1}{j} \\ &\times \frac{\theta}{1-q(1-\theta)}\left[\frac{p(1-\theta)}{1-q(1-\theta)}\right]^{x(n-k+j+1)+n+j-k} \end{aligned}$$

The $k^{th}$ order statistic's cdf is

$$\begin{aligned} F_k(x; p, \theta) &= \sum_{i=k}^{n}\binom{n}{i}\left[F\left(x; \lambda,\ \beta\right)\right]^i\left[1 - F\left(x; \lambda,\ \beta\right)\right]^{n-i} \\ &= \sum_{i=k}^{n}\sum_{j=0}^{n}(-1)^j\binom{n}{i}\binom{n}{j}\left[\frac{p(1-\theta)}{1-q(1-\theta)}\right]^{(x+1)(n-i+j)} \end{aligned}$$

Consider $X_{(1)} = min(X_1, X_2, \ldots, X_n)$, $X_{(n)} = max(X_1, X_2, \ldots, X_n)$, and $X_{(m+1)}$ with $m = \frac{n}{2}$ for minimum, maximum and the medium order statistics, respectively. As a result, the pmfs of the minimum, maximum, and median are

$$f_1\left(x; p,\ \beta\right) = \frac{n\theta}{1-q(1-\theta)}\left(\frac{p(1-\theta)}{1-q(1-\theta)}\right)^{n(x+1)-1}$$

$$\begin{aligned} f_n\left(x; p,\ \beta\right) &= \frac{n\theta}{1-q(1-\theta)}\left(\frac{p(1-\theta)}{1-q(1-\theta)}\right)^{x} \\ &\times \left[1 - \left(\frac{p(1-\theta)}{1-q(1-\theta)}\right)^{(x+1)}\right]^{n-1} \end{aligned}$$

$$\begin{aligned} f_{m+1}\left(x; p,\ \beta\right) &= \frac{n!}{(m)!\,(n-m+1)!}\frac{n\theta}{1-q\left(1-\theta\right)} \\ &\times \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{(n-m)(x+1)-1} \\ &\times \left[1 - \left(\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right)^{(x+1)}\right]^{m} \end{aligned}$$

### 3.4 Stress-strength reliability measure

One statistical measure that is frequently used to examine how well a system or material performs in terms of mechanical reliability is the stress-strength parameter. This measure denoted by the quantity $R = P(X > Y)$ where $Y$ represents the strength of the component and $X$ is the stress, while $R$ refers to the probability of the system performance and its complement $(1 - R)$ denotes the probability of the system failure. The stress-strength parameter is widely used in engineering, biostatistics, quality control, military, medicine and psychology.[16, 17] . In medicine, for the case control study the strength $Y$ represents outcomes of the treatment and the stress $X$ denotes the control group, while the quantity $R$ used to measure the ineffectiveness of the treatment. In many practical applications the stress-length model deals with a continuous quantitative data, so $X$ and $Y$ are considered to be independent continuous random variables.
In the discrete case, the stress-strength parameter is specified as

$$R = P\left(X > Y\right) = \sum_{x=0}^{\infty} f_X(x) F_Y(x)$$

where $f_X$ and $F_Y$ represent the pmf and cdf of the independent discrete random variables $X$ and $Y$, respectively.
Now, let $X \sim BG(p_1, \theta_1)$ and $Y \sim BG(p_2, \theta_2)$. Using Eqs.(4) and (5), we get

$$\begin{aligned} R &= \sum_{x=0}^{\infty}\left[\frac{\theta_1}{1-q_1(1-\theta_1)}\right]\left[\frac{p_1(1-\theta_1)}{1-q_1(1-\theta_1)}\right]^x \\ &\times \left\{1 - \left[\frac{p_2(1-\theta_2)}{1-q_2(1-\theta_2)}\right]^{x+1}\right\} \\ &= 1 - \frac{\theta_1 p_2(1-\theta_2)}{\left[1-q_1(1-\theta_1)\right]\left[1-q_2(1-\theta_2)\right] - p_1 p_2(1-\theta_1)(1-\theta_2)} \end{aligned}$$

### 3.5 Simulating the Random Sample

Random numbers from the BG distribution can be explored by equating the cdf of the distribution in Eq.(5) with a uniform random number and inverting the expression; that is, the random number from BG is obtained by solving the following Eq(12) for $x$.

$$1 - \left[\frac{p\left(1-\theta\right)}{1-q\left(1-\theta\right)}\right]^{x+1} = u \tag{12}$$

The random sample from BG can be further explained as

$$x = \frac{\log\left(1-u\right)}{\log\left[p\left(1-\theta\right)\right] - \log\left[1-q\left(1-\theta\right)\right]} - 1 \tag{13}$$

where $u$ is an arbitrary continuous uniform point within $(0, 1)$.

## 4 Maximum likelihood estimation

The purpose of this part is to determine the estimation of the BG distribution's parameters using the maximum likelihood estimation (MLE).
Let $X_1, X_2, \ldots, X_n$ represent a size $n$ random sample with a BG distribution. The log-likelihood is then

$$\begin{aligned} \ell &= \sum_{i=1}^{n}\log f(x) = \sum_{i=1}^{n}\log\left[\sum_{x=0}^{\infty}\frac{\theta\, p_i^x\left(1-\theta\right)^x}{\left[1-q\left(1-\theta\right)\right]^{x_i+1}}\right] \\ &= n\log\left[\frac{\theta}{1-(1-p)\left(1-\theta\right)}\right] \\ &\quad - \sum_{i=1}^{n} x_i \log\left[\frac{p\left(1-\theta\right)}{1-(1-p)\left(1-\theta\right)}\right] \end{aligned} \tag{14}$$

the likelihood equations are obtained by differentiating Eq.(14) partially with respect to the shape parameters $p$ and $\theta$ as

$$\frac{\partial \ell}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{(1-\theta)(\sum_{i=1}^{n} x_i + n)}{1-(1-p)(1-\theta)} \tag{15}$$

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} - \frac{(1-p)(\sum_{i=1}^{n} x_i + n)}{1-(1-p)(1-\theta)} - \frac{\sum_{i=1}^{n} x_i}{1-\theta} \tag{16}$$

Let $\frac{\partial \ell}{\partial p} = 0$ and $\frac{\partial \ell}{\partial \theta} = 0$ and solve the two equations to get MLEs $\hat{\tau} = (\hat{p}, \hat{\theta})'$ of $\tau = (p, \theta)'$. Since the MLE of the vector of unknown parameters $\tau = (p, \theta)'$ cannot be derived in closed forms, it is, therefore, hard to derive the exact distribution of the MLEs.

The second partial derivatives are given below

$$\frac{\partial^2 \ell}{\partial p^2} = \frac{(1-\theta)^2(\sum_{i=1}^{n} x_i + n)}{[1-(1-p)(1-\theta)]^2} - \frac{\sum_{i=1}^{n} x_i}{p^2}$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = \frac{(1-p)^2(\sum_{i=1}^{n} x_i + n)}{[1-(1-p)(1-\theta)]^2} - \frac{\sum_{i=1}^{n} x_i}{(1-\theta)^2} - \frac{n}{\theta^2}$$

$$\frac{\partial^2 \ell}{\partial p \partial \theta} = \frac{(1-p)(1-\theta)(\sum_{i=1}^{n} x_i + n)}{[1-(1-p)(1-\theta)]^2} + \frac{\sum_{i=1}^{n} x_i + n}{1-(1-p)(1-\theta)}$$

The asymptotic distribution of the MLE $\hat{\tau}$ is given as [18]

$$(\hat{\tau} - \tau) \to N\left[0,\ I^{-1}(\tau)\right]$$

where $I^{-1}(\tau)$ is the inverse of Fisher's information matrix of the unknown parameters $\tau = (p, \theta)'$ as follows:

$$I_{Y(p,\theta)}(\tau) = \begin{bmatrix} -E\left(\frac{\partial^2 \ell}{\partial p^2}\right) & -E\left(\frac{\partial^2 \ell}{\partial p \partial \theta}\right) \\ -E\left(\frac{\partial^2 \ell}{\partial p \partial \theta}\right) & -E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) \end{bmatrix}$$

On the other hand, Fisher's information matrix can be computed by using the approximation

$$I_Y(\hat{\tau}) = \begin{bmatrix} -\left.\frac{\partial^2 \ell}{\partial \theta^2}\right|_{(\hat{p},\hat{\theta})} & -\left.\frac{\partial^2 \ell}{\partial p \partial \theta}\right|_{(\hat{p},\hat{\theta})} \\ -\left.\frac{\partial^2 \ell}{\partial p \partial \theta}\right|_{(\hat{p},\hat{\theta})} & -\left.\frac{\partial^2 \ell}{\partial \theta^2}\right|_{(\hat{p},\hat{\theta})} \end{bmatrix}$$

where $\hat{p}$ and $\hat{\theta}$ are the MLEs of $p$ and $\theta$ respectively.

# 5 Simulation

This section demonstrates the Monte Carlo simulation that is conducted to examine the performance of the accuracy of point estimates for the parameters of proposed BG distribution using MLE with 10000 replications.

The simulation was performed with sample sizes of $n$ = 30, 100, 300 and 500 and some different values for the parameters $p$ and $\theta$. The empirical mean and the mean squared error (MSE) were used as criterions measures for evaluating the point estimation of the BG distribution's parameters. To generate the sample data from $BG(p, \theta)$ we used the formula of Eq.(13). Table 4 shows the empirical results of the simulation. All results in Table 4 explain that the estimates are close to the true values of the parameters for all sample sizes. Moreover, as the sample size increases, the MSEs decrease as expected, which means that the maximum likelihood method is appropriate for estimating the parameters of the BG distribution.

# 6 Application

In this section, we analyze two data sets to explain the applicability of the proposed BG distribution. Therefore, to examine the goodness of fit and to assess its performance, the BG distribution is compared with some related distributions including geometric and negative binomial (NB). The -log-likelihood (-log(L)) beside the $\chi^2$ statistic is used as criterions for comparison. BG distribution is fitted to two overdispersed right-skewed real lifetime count data sets from medical field. The first data reported by [19] (p12) explains the death times, in weeks of 30 patients with cancer of the tongue while second one shows the lengths of remission in weeks for a group of 30 leukemia patients receiving a particular kind of medication, this data set posted by [18](p219). Table 5 explains some descriptive measures for these two data sets. Looking at the results in Table 5, it is clear that both data sets are overdispersed right-skewed. Tables 6 and 7 demonstrate the MLE estimators, -LogL, and the $\chi^2$ statistics with its corresponding p-value for the geometric, NB as well as the proposed BG distribution for the two data sets tongue cancer and leukemia respectively. From the results in Tables 6 and 7, it can be seen that the values of -logL as well as the $\chi^2$ statistic for the three distributions almost equivalent. Moreover, the corresponding p-values indicate that the geometric, NB as well as the proposed BG distribution are appropriate for fitting both data sets.

# 7 Conclusion Remarks

This paper develops a new two shape parameters mixture of binomial distribution called the binomial-geometric (BG) distribution, created by combining the binomial with the geometric distribution using the probability generating function method. We look at some of the BG distribution statistical properties and use the maximum likelihood method to estimate its parameters. The proposed BG distribution is right-skewed with decreasing probability mass and constant hazard functions. The Shannon entropy, mean, and variance of the BG distribution increase with the parameter $p$, and decrease as the parameter $\theta$ increases. From the Monte Carlo simulation, we conclude that the maximum likelihood method can be used effectively to estimate the BG distribution parameters. Moreover, as $\theta$ increases, so do the coefficient of variation, skewness, and kurtosis, while these measures decrease as $p$ increases. On the other hand, the dispersion index grows up with $p$ and drops down as $\theta$ increases. The main finding that has been explored from the application results is that the BG distribution is the most convenient to fit both data sets that have been exampled. Therefore,

**Table 4.** Empirical means and MSEs for Simulation

| | | $p$=0.1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | $\theta$=0.01 | | $\theta$=0.05 | | $\theta$=0,1 | |
| $n$ | | $\hat{p}$ | $\hat{\theta}$ | $\hat{p}$ | $\hat{\theta}$ | $\hat{p}$ | $\hat{\theta}$ |
| 30 | Value | 0.1055 | 0.0109 | 0.1032 | 0.0531 | 0.1005 | 0.1034 |
| | MSE | 0.0226 | 0.000274 | 0.44822 | 0.118341 | 1.31947 | 1.16382 |
| 100 | Value | 0.1058 | 0.0107 | 0.1043 | 0.0525 | 0.1025 | 0.1028 |
| | MSE | 0.009574 | 0.000069 | 0.13839 | 0.032631 | 0.415486 | 0.341699 |
| 300 | Value | 0.1054 | 0.0106 | 0.1044 | 0.0523 | 0.1026 | 0.1026 |
| | MSE | 0.002163 | 0.000022 | 0.046195 | 0.010561 | 0.14071 | 0.11405 |
| 500 | Value | 0.1047 | 0.0105 | 0.1041 | 0.0521 | 0.1025 | 0.1024 |
| | MSE | 0.001272 | 0.000013 | 0.027396 | 0.006226 | 0.083983 | 0.067743 |
| | | $p$=0.5 | | | | | |
| | | $\theta$=0.01 | | $\theta$=0.05 | | $\theta$=0,1 | |
| $n$ | | $\hat{p}$ | $\hat{\theta}$ | $\hat{p}$ | $\hat{\theta}$ | $\hat{p}$ | $\hat{\theta}$ |
| 30 | Value | 0.5224 | 0.0107 | 0.5267 | 0.0542 | 0.5189 | 0.1062 |
| | MSE | 0.54577 | 0.000256 | 15.25184 | 0.167513 | 140.3161 | 8.18099 |
| 100 | Value | 0.5226 | 0.0105 | 0.5263 | 0.0529 | 0.5194 | 0.1042 |
| | MSE | 0.155025 | 0.000064 | 4.305085 | 0.04082 | 17.17046 | 0.571032 |
| 300 | Value | 0.5209 | 0.0104 | 0.5214 | 0.0522 | 0.5179 | 0.1035 |
| | MSE | 0.050923 | 0.00002 | 1.35872 | 0.012402 | 5.799814 | 0.188427 |
| 500 | Value | 0.5192 | 0.0104 | 0.5186 | 0.0519 | 0.5161 | 0.1032 |
| | MSE | 0.030241 | 0.000012 | 0.798734 | 0.007258 | 3.375126 | 0.109206 |
| | | $p$=0.9 | | | | | |
| | | $\theta$=0.01 | | $\theta$=0.05 | | $\theta$=0,1 | |
| $n$ | | $\hat{p}$ | $\hat{\theta}$ | $\hat{p}$ | $\hat{\theta}$ | $\hat{p}$ | $\hat{\theta}$ |
| 30 | Value | 0.9505 | 0.0109 | 0.9403 | 0.0534 | 0.9471 | 0.1076 |
| | MSE | 1.872064 | 0.000283 | 47.45927 | 0.157115 | 229.7787 | 2.60926 |
| 100 | Value | 0.9501 | 0.0106 | 0.9375 | 0.0524 | 0.9493 | 0.1057 |
| | MSE | 0.530355 | 0.000069 | 13.48123 | 0.039358 | 65.36218 | 0.671546 |
| 300 | Value | 0.9422 | 0.0105 | 0.9332 | 0.0519 | 0.9453 | 0.1048 |
| | MSE | 0.16929 | 0.000021 | 4.34354 | 0.012203 | 21.76659 | 0.217836 |
| 500 | Value | 0.9377 | 0.0104 | 0.931 | 0.0517 | 0.9404 | 0.1042 |
| | MSE | 0.099583 | 0.000012 | 2.569032 | 0.007167 | 14.47637 | 0.145657 |

**Table 5.** descriptive measures for the two data sets

| Data set | $n$ | Min | max | Mean | Variance | Skewness |
|---|---|---|---|---|---|---|
| Tongue cancer | 30 | 1 | 167 | 50.0312 | 1945.838 | 0.97211 |
| Leukemia | 30 | 1 | 91 | 25.3333 | 662.092 | 1.192072 |

**Table 6.** estimates of the parameters, -log (L), k-s test value and p-value for the chosen distributions of the tongue cancer patient's data set

| Model | Parameters estimates | | -log(L) | k-s | P-value |
|---|---|---|---|---|---|
| BG | $\hat{p} = 0.1236$ | $\hat{\theta} = 002464$ | 157.5224 | 23.69 | 0.5935 |
| Geometric | | $\hat{\theta} = 0.01961$ | 157.5224 | 23.75 | 0.5900 |
| NB | $\hat{p} = 0.01965$ | $\hat{r} = 1.0025$ | 157.5224 | 23.71 | 0.5922 |

**Table 7.** estimates of the parameters, -log (L), k-s test value and p-value for the chosen distributions of the tongue leukemia patient's data set

| Model | Parameters estimates | | -log(L) | k-s | P-value |
|---|---|---|---|---|---|
| BG | $\hat{p} = 0.1128$ | $\hat{\theta} = 00443$ | 127.54 | 33.05 | 0.1030 |
| Geometric | | $\hat{\theta} = 0.3798$ | 127.54 | 33.06 | 0.1026 |
| NB | $\hat{p} = 0.03907$ | $\hat{r} = 1.02983$ | 127.54 | 33.42 | 0.0954 |

we recommend the proposed BG distribution for modelling the overdispersed right-skewed real-life count data sets adequacy because the two parameters make BG distribution more flexible to suit most types of count data sets. Further, the proposed BG distribution can be an alternative to the negative binomial and geometric distributions.

## Acknowledgements

We're very thankful for the editor's and anonymous reviewers' helpful feedback and suggestions, which made this work much better.

## REFERENCES

[1] W. R. Blischke. Estimating the parameters of mixtures of binomial distributions, Journal of the American Statistical Association, Vol.59, No.306, 510-528, 1964.

[2] D. C. Schmittlein. Surprising inferences from unsurprising observations: do conditional expectations really regress to the mean?, The American Statistician, Vol.43, No.3, 176-183, 1989.

[3] M. K. Roy, S. Rahman, M. M. Ali. A class of poisson mixtured distributions. Journal of Information and Optimization sciences, Vol.13, No.2, 207-218, 1992.

[4] G. Wood. Binomial mixtures and Finite Exchangeability, The Annals of Statistics, Vol.20, No.3, 1992, DOI: 10.1214/aop/1176989684

[5] G. Wood. Binomial mixtures: Geometric estimation of the mixing distribution, The Annals of Statistics, Vol.27, No.5, 1999, DOI: 10.1214/aos/1017939148

[6] M. K. Roy, A. K. Roy, M. M. Ali. Binomial Mixtures of some Standard Distributions, Journal of Information and Optimization Sciences, Vol.14, No.1, 57-71, 1993, DOI: 10.1080/02522667.1993.10699136

[7] J. Zhu, J. C. Eickhoff, M. S. Kaiser. Modeling the Dependence between Number of Trials and Success Probability in Beta-Binomial-Poisson Mixture Distributions, Biometrics, Vol. 59, No. 4, 955-961, 2003, DOI: 10.1111/j.0006-341X.2003.00110.x

[8] M. J. Faddy, D. M. Smith. Modeling the Dependence between the Number of Trials and the Success Probability in Binary Trials. Biometrics, Vol. 61, N0. 4, 1112-1114, 2005.

[9] Z. Shkedy, G. Molenberghs, H. V. Craenendonck, T. Steckler, L. Bijnens. A hierarchical Binomial-Poisson model for the analysis of a crossover design for correlated binary data when the number of trials is dose-dependent, Journal of Biopharmaceutical Statistics, Vol. 15, N0. 2, 225-239, 2005, DOI: 10.1081/BIP-200049825

[10] L. Grilli, C. Rampichini, R. Varriale. Binomial Mixture Modeling of University Credits. Communication in Statistics-Theory and Methods, Vol. 44, N0. 22, 4866-4879, 2015, DOI: 10.1080/03610926.2013.804565

[11] J. Knape, D. Arlt, F. Barraquand, A. Berg, M. Chevalier, T. Pärt, A. Ruete, M. Żmihorski. Sensitivity of binomial N-mixture models to overdispersion: The importance of assessing model fit. Methods in Ecology and Evolution, Vol. 9,2102–2114, 2018, DOI: 10.1111/2041-210X.13062

[12] A. R. El-Alosey. Random Sum of New Type of Mixtures of Distributions. International Journal of Statistics and Systems, Vol. 2, N0. 1, 49-57, 2007.

[13] K. Abed Al-Kadim, R. N. AL-Hussani. Binomial mixture of Erlang distribution. International Journal of Mathematics and Statistics Studies, Vol. 4, N0. 2, 28-38, 2016.

[14] M. A. S. Adnan, H. Kiser. A class of triple mixture distributions. Far East Journal of Theoretical Statistics, Vol. 59, N0. 2, 59-79, 2020, http://dx.doi.org/10.17654/TS059020059

[15] Adnan M. A. S., H. Kiser, A. S. Adnan, S. Shamsi. A class of power function mixture distributions. Far East Journal of Theoretical Statistics, Vol. 61, N0. 2, 191-208, 2021, http://dx.doi.org/10.17654/TS061020191

[16] S. Kotz, M. Lumelskii, M. Pensky. The Stress-Strength Model and its Generalizations: Theory and Applications. Work Scientific, New-York, 2003.

[17] L. Ventura, W. Racugno. Recent advances on Bayesian inference for $P(X < Y)$. Bayesian analysis, Vol. 6, N0. 2:1–75, 2011.

[18] J. F. Lawless. Statistical models and methods for lifetime data, John Wiley and Sons, 2011.

[19] J. P. Klein, M. L. Moeschberger. Survival analysis: techniques for censored and truncated data, Vol. 2, pp. 3-5, New York: Springer, 2003.