

An Extension of the Geometric Distribution with Properties and Applications

Subrata Chakraborty
Dibrugarh University

Seng Huat Ong
UCSI University

Aniket Biswas
Dibrugarh University

Abstract

A new parameter is introduced to extend the geometric distribution using Azzalini's method. Several important structural properties of the proposed two-parameter extended geometric distribution are investigated. Characterizations including for the geometric distribution, in terms of the proposed model, are established. Maximum likelihood estimation, method of moment estimation and relative frequency based estimation of the parameters are discussed in detail. The likelihood ratio test regarding relevance of the additional parameter is presented. Bayesian estimation of the parameters using STAN is also discussed. The proposed model is compared with some recently introduced two-parameter count models by analyzing two real-life datasets. The findings clearly indicate superiority of the proposed model over the rest.

Keywords: skewed distribution, increasing failure rate, log-concave, unimodal, recurrence relation.

1. Introduction

There are several existing methods of generating and generalizing continuous probability distributions. In particular, skewed continuous distributions have been thoroughly examined and extended by many researchers stemming from the popular skew normal distribution of [Azzalini \(1985\)](#). The standard skew normal distribution with skewing parameter $\alpha \in (-\infty, \infty)$ has the probability density function (pdf)

$$f(z) = 2\phi(z)\Phi(\alpha z), \quad z \in (-\infty, \infty). \quad (1)$$

Here, ϕ and Φ denote the pdf and cumulative distribution function (cdf) of the standard normal distribution. Let g and G denote pdf and cdf of a continuous distribution with support $\mathcal{S} \subset (-\infty, \infty)$, respectively. Thus in general,

$$f(z) = \frac{g(z)G(\alpha z)}{\int_{\mathcal{S}} g(t)G(\alpha t)dt}, \quad z \in \mathcal{S}. \quad (2)$$

Clearly, Azzalini's method is a particular case of the above weighting mechanism with cdf as the weighting function. While numerous continuous distributions are introduced in last three decades with Azzalini's method, such attempt with discrete distributions is unavailable.

Thus, it is of interest to apply Azzalini's method to a discrete distribution with compact cdf, such as the geometric distribution. Extension of geometric distribution has attracted many researchers over the years. For instance, Jain and Consul (1971), Philippou, Georgiou, and Philippou (1983), Tripathi, Gupta, and White (1987), Makcutek (2008), Gómez-Déniz (2010) and Nekoukhou, Alamatsaz, and Bidram (2012). Many of these generalizations rely on the discretization of continuous distributions or extension using the Marshall-Olkin family of distributions. Recently, Chakraborty and Gupta (2015) proposed a generalization known as the exponentiated geometric distribution and Bhati, Sastry, and Maha Qadri (2015) introduced a weighted geometric distribution. The extended geometric distribution to be introduced in this article by using Azzalini's method, is different from the generalizations of the geometric distribution mentioned above. It is to be noted that the motivation behind the use of Azzalini's method, is for extending the geometric distribution, unlike skewing a symmetric distribution.

In the next section, we introduce the extended geometric distribution. We investigate its distributional properties in Section 3. In Section 4, we discuss a few related characterizations. Section 5 is about estimation of the parameters in the model from both the frequentist and Bayesian perspectives. Two real life applications of the proposed distribution are presented in Section 6. We conclude the article by pointing out several scopes and some limitations of the current work.

2. Proposed model

Consider the random experiment of tossing a coin independently until head appears. Let the probability of obtaining head in each toss be p , $0 < p < 1$. If Y denotes the number of tails obtained before getting the first head, then Y has the following probability mass function (pmf).

$$p_Y(y) = pq^y, \quad q = 1 - p, \quad y = 0, 1, 2, \dots \quad (3)$$

Here, Y is said to follow the geometric distribution with parameter p and we write $Y \sim G(p)$. The cdf of Y is given below.

$$F_Y(y) = 1 - q^{y+1} \quad (4)$$

Since, Y has non-negative support, $F_Y(\alpha y) = 0$ for $\alpha < 0$. In analogy with (2), we propose an extension of the geometric distribution with the following pmf.

$$p_X(x) = \frac{p_Y(x)F_Y(\alpha x)}{\sum_{x=0}^{\infty} p_Y(x)F_Y(\alpha x)} = \frac{pq^x(1 - q^{\alpha x+1})}{W(p, \alpha)}. \quad (5)$$

Here, $W(p, \alpha) = (1 - q^{\alpha+1} - pq)/(1 - q^{\alpha+1})$, $0 < p < 1$, $\alpha > 0$ and $x = 0, 1, 2, \dots$. The distribution of X is named the extended geometric distribution and we write $X \sim EG(p, \alpha)$.

Special cases:

- $EG(p, 0)$ reduces to $G(p)$.
- For $\alpha \rightarrow \infty$,

$$p_X(x) = \begin{cases} \frac{pq^x}{1-pq}, & x = 1, 2, \dots \\ \frac{p^2}{1-pq}, & x = 0. \end{cases}$$

Remark 1. For large α , $EG(p, \alpha)$ can accommodate excess zeroes for $p \rightarrow 1$. However, for $p \rightarrow 0$, $EG(p, \alpha)$ turns out to be zero-deflated.

Now, we present a stochastic formulation of the $EG(p, \alpha)$ distribution considering the following coin tossing experiment. Consider two persons A and B tossing two different coins

independently. Both the coins have same probability p of showing head. Let the number of tails obtained by A and B before the first head be U and V , respectively. If $V \leq U$ is known, then U is distributed as $EG(p, 1)$. Generalizing this situation with an additional parameter $\alpha \in \{0, 1, 2, \dots\}$, $EG(p, \alpha)$ can be viewed as the distribution of U when it is known that $V \leq \alpha U$. The genesis of $EG(p, 0)$ or $G(p)$ is clear when we put $\alpha = 0$ in the conditional statement $V \leq \alpha U$. This formulation can be justified by the following. Clearly, U and V are iid $G(p)$. Then

$$\begin{aligned} P(U = u | V \leq \alpha U) &= \frac{P(\{U = u\} \cap \{V \leq \alpha U\})}{P(V \leq \alpha U)} \\ &= \frac{P(U = u)P(V \leq \alpha u)}{P(V \leq \alpha U)} \\ &= \frac{pq^u(1 - q^{\alpha u + 1})}{\sum_{j=0}^{\infty} F_X(\alpha j)p_X(j)} \\ &= \frac{pq^u(1 - q^{\alpha u + 1})}{W(p, \alpha)}. \end{aligned}$$

The converse is also true and we discuss it further in Section 4. Though the formulation given above is for non-negative integer α , we study the proposed distribution with non-negative real α as motivated from Azzalini's mechanism. Remark 1 is justified by Figure 1. For larger values of α and $p = 0.5$, the distribution has two modes, at 0 and 1. For smaller values of p , the shape of the distribution changes from the J-shaped geometric distribution with non-zero mode. For larger values of p , the distribution is able to accommodate outliers. The distribution has longer tail for smaller values of p .

3. Distributional properties

The $EG(p, \alpha)$ model, proposed in the last section, is simple, having only two parameters, flexible and has some useful shapes. Therefore in this section, we explore different properties of the proposed model in detail.

3.1. Log-concavity and unimodality

A discrete distribution is said to be log-concave if its pmf f satisfies $f^2(k) \geq f(k+1)f(k-1)$. Note that, for $k = 1, 2, \dots$

$$\begin{aligned} \{pq^k(1 - q^{\alpha k + 1})\}^2 &\geq pq^{k+1}(1 - q^{\alpha(k+1)+1})pq^{k-1}(1 - q^{\alpha(k-1)+1}) \\ \implies p_X^2(k) &\geq p_X(k+1)p_X(k-1). \end{aligned}$$

Thus, $EG(p, \alpha)$ is log-concave and hence by Theorem 3 of [Keilson and Gerber \(1971\)](#), it is also strongly unimodal.

Remark 2. The $EG(p, \alpha)$ distribution has non-zero mode if $\alpha > \log(2q - 1)/\log(q) - 2$. For example, the distribution has non-zero mode for $p = 0.25$ and $\alpha > 0.409$. In case, when $\alpha = \log(2q - 1)/\log(q) - 2$, we have two modes at $x = 0$ and 1. For $p = 0.25$ and $\alpha = 0.409$, unimodality, in the sense of having a unique modal region defined by the set $\{0, 1\}$, having the same mass equal to the maximum. That is, $p_X(0) = p_X(1) = \max_{x \in \{0, 1, \dots\}} p_X(x)$.

3.2. Recurrence relation and tail length

Probability recurrence relation plays an important role in computing mass function and in determining tail behaviour of a distribution. For fixed x ,

$$\frac{p_X(x+1)}{p_X(x)} = \frac{q(1 - q^{\alpha x + \alpha + 1})}{1 - q^{\alpha x + 1}}. \quad (6)$$

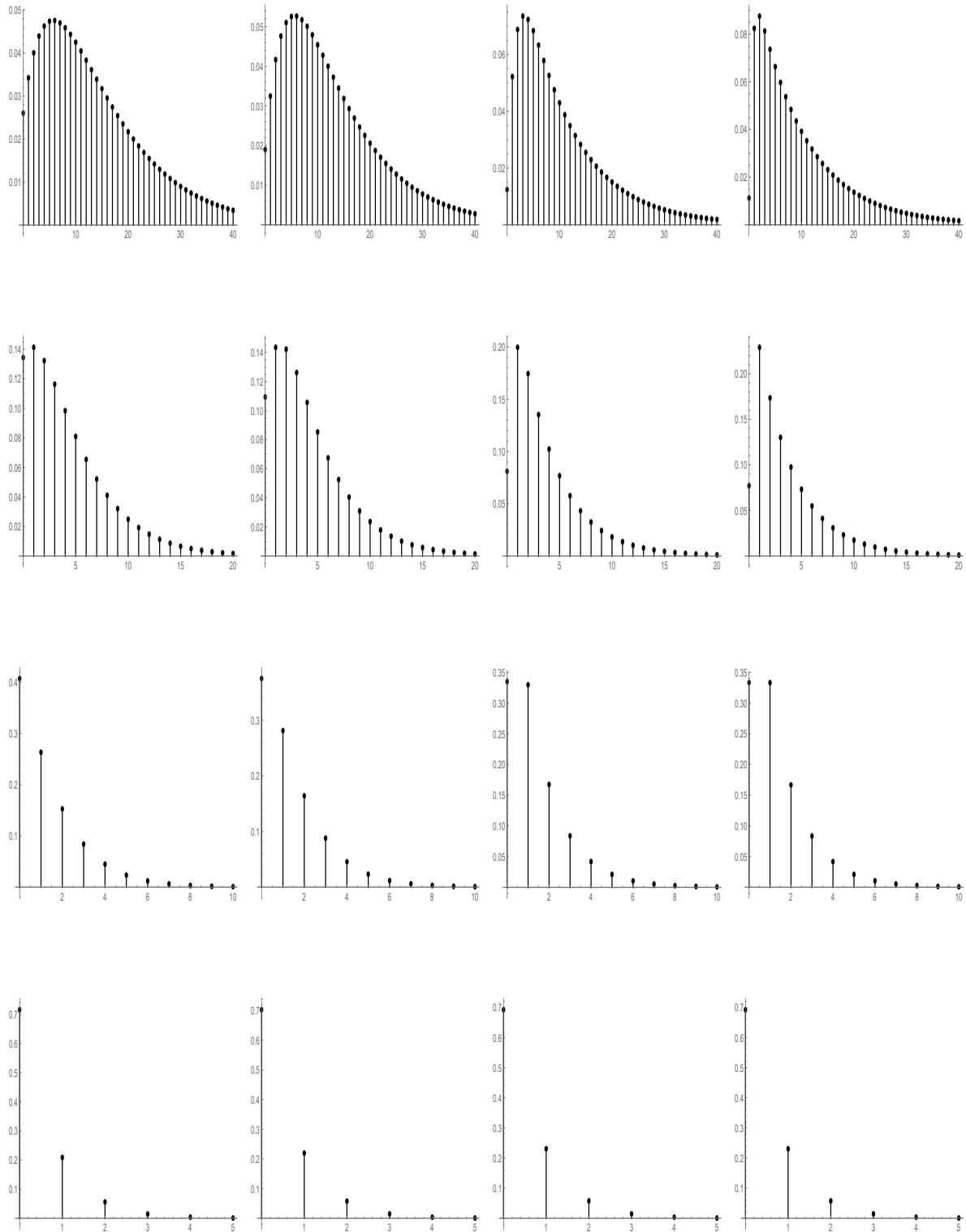


Figure 1: Probability mass function of $EG(p, \alpha)$ for $p \in \{0.1, 0.25, 0.50, 0.75\}$ and $\alpha \in \{0.5, 1, 5, 15\}$. The (i, j) -th plot corresponds to i -th value of p and j -th value of α for $i, j = 1, 2, 3, 4$.

From (6), we have the following two-term recurrence formula.

$$p_X(x+1) = \frac{q(1 - q^{\alpha x + \alpha + 1})}{1 - q^{\alpha x + 1}} p_X(x). \quad (7)$$

The behaviour of the tail of the distribution may be studied from the ratio of probabilities (Ong and Muthaloo 2012). As $x \rightarrow \infty$ the RHS of (6) tends to q . Hence the distribution is short or long tailed depending on the value of p . When $p \rightarrow 1$, we get a Poisson-type tail and for $p \rightarrow 0$, a longer tail appears.

3.3. Cumulative distribution function

Using (5) in $F_X(x) = \sum_{j=1}^x p_X(j)$, we have

$$\begin{aligned} F_X(x) &= \frac{1}{W(p, \alpha)} \sum_{j=0}^x pq^j (1 - q^{\alpha j + 1}) \\ &= \frac{1}{W(p, \alpha)} \left[\sum_{j=0}^x pq^j - \sum_{j=0}^x pq^{\alpha j + j + 1} \right] \\ &= \frac{1}{W(p, \alpha)} \left[(1 - q^{x+1}) - pq \frac{1 - q^{(\alpha+1)(x+1)}}{1 - q^{\alpha+1}} \right]. \end{aligned} \quad (8)$$

3.4. Random sample generation

As an explicit form of the cdf is available, one can use probability integral transform to generate a random observation x from $EG(p, \alpha)$ by solving $F_X(x) = u$ for x , given the values of p, α and $u \sim \text{Uniform}(0, 1)$. Note that, $F_X(x) = u$ implies

$$1 - uW(p, \alpha) = q^{x+1} + C[1 - q^{(\alpha+1)(x+1)}]. \quad (9)$$

Here, $C = pq/(1 - q^{\alpha+1})$. After some algebraic patchwork, from (9), we obtain

$$B = y[1 - Cy^\alpha]. \quad (10)$$

Here, $B = 1 - C - uW(p, \alpha)$ and $y = q^{x+1}$. It is obvious from (10), for the proposed distribution, a direct solution for x , as outlined above, is not feasible. Application of different numerical routines, for solving (10), yield inconsistent and imprecise results. Note that, $y \in (0, 1)$ since $q \in (0, 1)$. Thus, we prescribe the following exhaustive search based algorithm that provides consistently precise output.

Algorithm 1

(For generating a sample of size n from $EG(p, \alpha)$)

- *Step-1:* Input p and α .
- *Step-2:* Compute: $W = W(p, \alpha)$, $C = pq/(1 - q^{\alpha+1})$.
- *Step-3:* Generate one u from $\text{Uniform}(0, 1)$ distribution and compute $B = 1 - C - uW$.
- *Step-4:* For $y = \{0.0001, 0.0002, \dots, 0.9999\}$, obtain $y_0 = \min_y \{y - Cy^{\alpha+1} - B\}$.
- *Step-5:* Find $x = \lfloor \log(y_0)/\log(q) \rfloor$.

- *Step-6*: Repeat *Step-2* to *Step-5* n times.

3.5. Generating functions

Using (5) in $P_X(t) = \sum_{j=0}^{\infty} t^j p_X(j)$, we get the probability generating function (pgf) as

$$\begin{aligned} P_X(t) &= \frac{1}{W(p, \alpha)} \sum_{j=0}^{\infty} p(tq)^j (1 - q^{\alpha j + 1}) \\ &= \frac{1}{W(p, \alpha)} [H(t) - qH(q^{\alpha}t)]. \end{aligned} \quad (11)$$

Here, $H(t) = p/(1 - qt)$ is the pgf of $G(p)$. Similarly, the moment generating function (mgf) of X is readily obtained in terms of $Q(t)$, the mgf of $G(p)$, as follows

$$M_X(t) = \frac{1}{W(p, \alpha)} [Q(t) - qQ(q^{\alpha}t)]. \quad (12)$$

3.6. Moments

The r^{th} factorial moment can be obtained by differentiating the pgf in (11) r times with respect to t and putting $t = 1$.

Now,

$$\begin{aligned} \frac{\delta^r}{\delta t^r} H(t)|_{t=1} &= r! \left(\frac{q}{p}\right)^r \\ \frac{\delta^r}{\delta t^r} H(q^{\alpha}t)|_{t=1} &= \frac{r! pq^{(\alpha+1)r}}{(1 - q^{\alpha+1})^{r+1}}. \end{aligned}$$

Therefore using (11), we get

$$E(X_{(r)}) = \frac{r!}{W(p, \alpha)} \left[\left(\frac{q}{p}\right)^r - \frac{pq^{(\alpha+1)r+1}}{(1 - q^{\alpha+1})^{r+1}} \right]$$

Where $x_{(r)} = x(x-1)\dots(x-r+1)$. Hence the first two raw moments of $EG(p, \alpha)$ are obtained below.

$$\begin{aligned} E(X) &= \frac{1}{W(p, \alpha)} \left[\frac{q}{p} - \frac{pq^{(\alpha+2)}}{(1 - q^{\alpha+1})^2} \right] \\ &= \frac{q\{1 - 3q^{\alpha+1} + 2q^{\alpha+2} - q^{\alpha+3} + q^{2\alpha+2}\}}{p(1 - q^{\alpha+1})(1 - q + q^2 - q^{\alpha+1})} \end{aligned} \quad (13)$$

$$E(X_{(2)}) = \frac{2}{W(p, \alpha)} \left[\left(\frac{q}{p}\right)^2 - \frac{pq^{2\alpha+3}}{(1 - q^{\alpha+1})^3} \right] \quad (14)$$

Variance follows immediately as

$$\begin{aligned} Var(X) &= E(X_{(2)}) + E(X) - (E(X))^2 \\ &= \frac{q}{p^2} + \frac{q^{\alpha+1}}{(1 - q^{\alpha+1})^2 - \frac{2p+q^2-q^{\alpha+1}}{p+q^2-q^{\alpha+1}}}. \end{aligned} \quad (15)$$

Other higher order moments can be derived similarly. The index of dispersion (ID) is given by

$$\begin{aligned} ID_X &= \frac{Var(X)}{E(X)} \\ &= \frac{1}{p} + \frac{1}{1 - q^{\alpha+1}} + \frac{2\{(q-2)q^{\alpha+1} - 1\}}{1 + q^{2(\alpha+1)} - q^{\alpha+1}\{q(q-2) + 3\}} - \frac{p}{q(q^{\alpha} - q + 1)} \end{aligned}$$

From Figure 2, the distribution is over dispersed and the ID can be very high for small values of p .

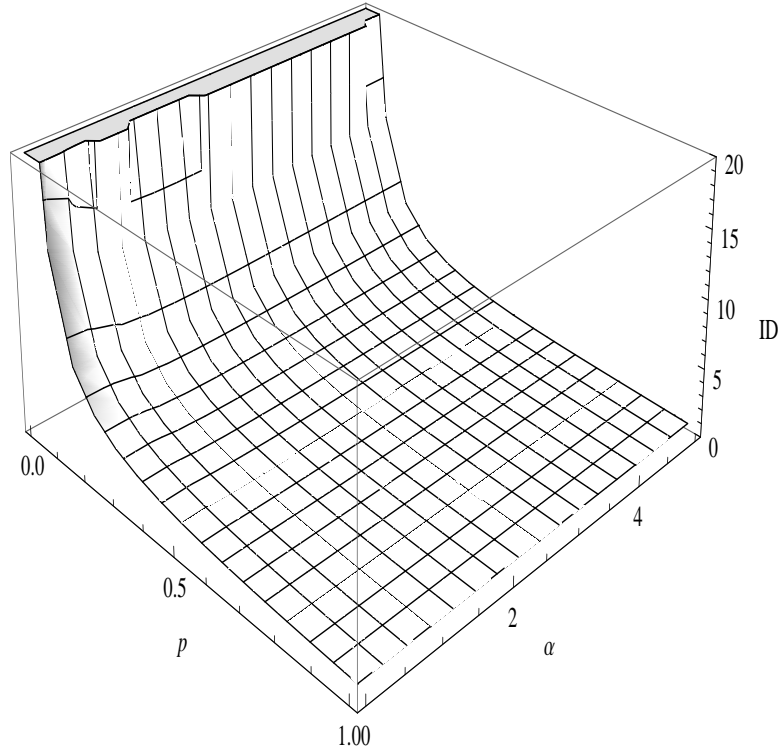


Figure 2: Index of dispersion of $EG(p, \alpha)$ for $p \in (0, 1)$ and $\alpha \in (0, 5)$

3.7. Reliability properties

The reliability function of a discrete random variable T at time t is the probability that the survival time T is at least t . From (8), the reliability function of $EG(p, \alpha)$ is given by

$$R_X(x) = P(X \geq x) = 1 - \frac{1}{W(p, \alpha)} \left[(1 - q^x) - pq \frac{1 - q^{(\alpha+1)x}}{1 - q^{\alpha+1}} \right]. \quad (16)$$

The hazard function of a discrete random variable T at time point t is defined as the conditional probability of failure at t , given that the survival time T is at least t . Now, using (5) and (16), the hazard function is obtained as follows.

$$h_X(x) = \frac{P(X = x)}{P(X \geq x)} = \frac{p_X(x)}{R_X(x)} = \frac{p(1 - q^{\alpha x+1})(1 - q^{\alpha+1})}{1 - q^{\alpha+1} - pq^{\alpha x+1}} \quad (17)$$

From Figure 3, it is seen that $EG(p, \alpha)$ has monotonically increasing hazard rate.

4. Characterizations

We investigate three interesting stochastic characterizations of the proposed extension of the geometric distribution through the following results.

Theorem 1. *Let U and V be independently and identically distributed (iid) random variables with support $\{0, 1, 2, \dots\}$. For any $\alpha \in \{0, 1, 2, \dots\}$, the conditional distribution of $U|V \leq \alpha U$ follows $EG(p, \alpha)$ if and only if (iff) $U, V \sim G(p)$.*

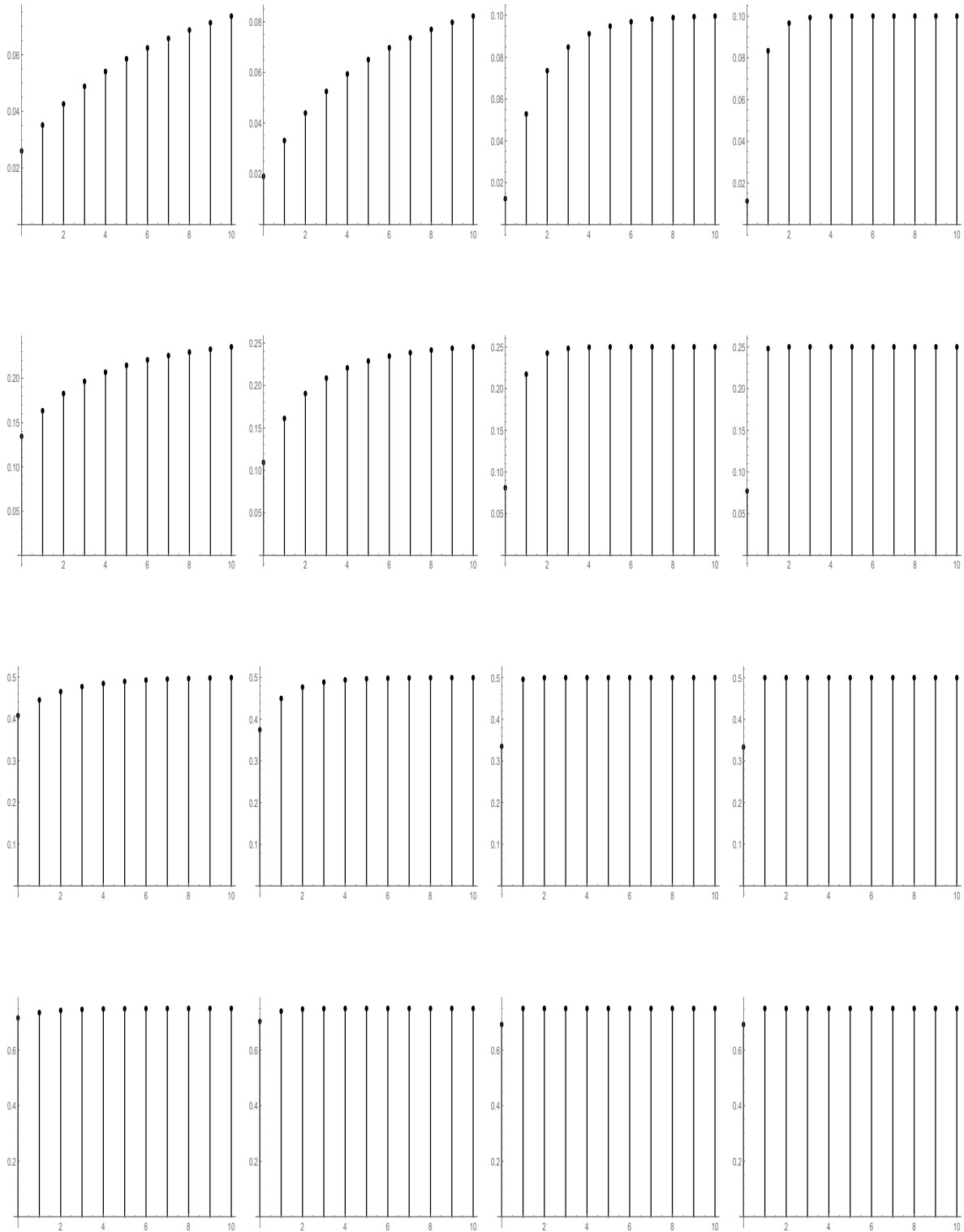


Figure 3: Hazard function of $EG(p, \alpha)$ for $p \in \{0.1, 0.25, 0.50, 0.75\}$ and $\alpha \in \{0.5, 1, 5, 15\}$. The (i, j) -th plot corresponds to i -th value of p and j -th value of α for $i, j = 1, 2, 3, 4$.

Proof. If part: Discussed in Section 2.

Only if part: Note that,

$$\begin{aligned} P[U = u|V \leq \alpha U] &= \frac{pq^u((1 - q^{\alpha u+1})(1 - q^{\alpha+1}))}{1 - q^{\alpha+1} - pq} \\ \implies \frac{P(U = u)F_V(\alpha u)}{P(V \leq \alpha U)} &= \frac{pq^u((1 - q^{\alpha u+1})(1 - q^{\alpha+1}))}{1 - q^{\alpha+1} - pq}. \end{aligned}$$

In particular, for $\alpha = 1$,

$$\begin{aligned} P(U = u)F_V(u) &= P(V \leq U) \frac{pq^u((1 - q^{\alpha u+1})(1 - q^{\alpha+1}))}{1 - q^{\alpha+1} - pq} \\ &= Cq^u(1 - q^{u+1}). \end{aligned} \quad (18)$$

Here, C is some constant, independent of u . Putting $u = 0$ in (18), we get

$$\begin{aligned} P(U = 0)P(V = 0) &= C(1 - q) \\ \implies [P(U = 0)]^2 &= C(1 - q). \end{aligned} \quad (19)$$

Similarly, for $u = 1$,

$$\begin{aligned} P(U = 1)F_V(1) &= Cq(1 - q^2) \\ \implies P(U = 1)P(U = 0) + [P(U = 1)]^2 &= Cq(1 - q^2). \end{aligned} \quad (20)$$

Let, $\theta_k = P(U = k)/P(U = 0)$ for $k = 1, 2, \dots$ we get

$$\begin{aligned} \theta_1[P(U = 0)]^2 + \theta_1^2[P(U = 0)]^2 &= Cq(1 - q^2) \\ \implies \theta_1^2 + \theta_1 - q(1 + q) &= 0 \quad (\text{Using (19)}) \\ \implies \theta_1 = q \quad \text{or} \quad -1 - q. \end{aligned} \quad (21)$$

Hence,

$$P(U = 1) = qP(U = 0). \quad (22)$$

Now, we claim that,

$$P(U = u) = q^u P(U = 0). \quad (23)$$

We prove (23) by induction. Note that, (23) is true for $u = 1$, by (22). Let (23) be true $u = n$. Now,

$$\begin{aligned} P(U = n+1)F_V(n+1) &= Cq^{n+1}(1 - q^{n+2}) \quad \text{From (18)} \\ \implies P(U = n+1)\{P(U = 0) + \dots + P(U = n+1)\} &= Cq^{n+1}(1 - q^{n+2}) \\ \implies \theta_{n+1}P(U = 0)\{P(U = 0) + \dots + q^n P(U = 0) + P(U = n+1)\} &= Cq^{n+1}(1 - q^{n+2}) \\ \implies \theta_{n+1}[P(U = 0)]^2\{1 + q + \dots + q^n + \theta_{n+1}\} &= Cq^{n+1}(1 - q^{n+2}) \\ \implies \theta_{n+1}^2 + \theta_{n+1} \sum_{j=0}^n q^j - q^{n+1} \sum_{j=0}^{n+1} q^j &= 0 \quad \text{Using (19)} \\ \implies \theta_{n+1} = q^{n+1} \quad \text{or} \quad - \sum_{j=0}^{n+1} q^j. \end{aligned} \quad (24)$$

Hence,

$$\begin{aligned} \theta_{n+1} &= \frac{P(U = n+1)}{P(U = 0)} = q^{n+1} \\ \implies P(U = n+1) &= q^{n+1} P(U = 0). \end{aligned} \quad (25)$$

Therefore $P(U = u) = q^u P(U = 0) = pq^u$ and thus, $U \sim G(p)$ and U, V being identical $V \sim G(p)$. \square

Theorem 2. Let X be any discrete random variable with support $\{0, 1, 2, \dots\}$. For any $\alpha > 0$, X follows $EG(p, \alpha)$ iff

$$E \left[q^X (1+q)(1-q^{\alpha+1}) - pq^{X(\alpha+1)+1} | X \geq k \right] = q^k (1 - q^{\alpha+1} - pq^{\alpha k+1}). \quad (26)$$

Proof. If part

We have seen that (26) can be written as

$$\begin{aligned} \sum_{x \geq k} \left[q^x (1 - q^{\alpha+1} - pq^{\alpha x+1}) + q^{x+1} (1 - q^{\alpha+1} - pq^{\alpha(x+1)+1}) \right] & \frac{P(X=x)}{P(X \geq k)} \\ & = q^k (1 - q^{\alpha+1} - pq^{\alpha k+1}). \end{aligned} \quad (27)$$

From this we write

$$\begin{aligned} \sum_{x \geq k+1} \left[q^x (1 - q^{\alpha+1} - pq^{\alpha x+1}) + q^{x+1} (1 - q^{\alpha+1} - pq^{\alpha(x+1)+1}) \right] & \frac{P(X=x)}{P(X \geq k+1)} \\ & = q^{k+1} (1 - q^{\alpha+1} - pq^{\alpha(k+1)+1}). \end{aligned} \quad (28)$$

Taking difference of (27) and (28) we get

$$\begin{aligned} & \left[q^k (1 - q^{\alpha+1} - pq^{\alpha k+1}) + q^{k+1} (1 - q^{\alpha+1} - pq^{\alpha(k+1)+1}) \right] P(X=k) \\ & = \frac{(q^k (1 - q^{\alpha+1} - pq^{\alpha k+1}))^2 - (q^{k+1} (1 - q^{\alpha+1} - pq^{\alpha(k+1)+1}))^2}{1 - q^{\alpha+1} - pq} \\ \implies P(X=k) & = \frac{(1-q)(1-q^{\alpha k+1})(1-q^{\alpha+1})}{1 - q^{\alpha+1} - pq} \end{aligned}$$

Only if part

$$\begin{aligned} & E \left[q^X (1+q)(1-q^{\alpha+1}) - pq^{X(\alpha+1)+1} (1+q^{\alpha+1}) | X \geq k \right] \\ & = \sum_{x \geq k} \left[q^x (1 - q^{\alpha+1} - pq^{\alpha x+1}) + q^{x+1} (1 - q^{\alpha+1} - pq^{\alpha(x+1)+1}) \right] \frac{P(X=x)}{P(X \geq k)} \\ & = \frac{1}{P(X \geq k)} \sum_{x \geq k} \left[q^x (1 - q^{\alpha+1} - pq^{\alpha x+1}) + q^{x+1} (1 - q^{\alpha+1} - pq^{\alpha(x+1)+1}) \right] P(X=x) \\ & = \frac{1}{P(X \geq k)} \sum_{x \geq k} \frac{(q^x (1 - q^{\alpha+1} - pq^{\alpha x+1}))^2 - (q^{x+1} (1 - q^{\alpha+1} - pq^{\alpha(x+1)+1}))^2}{q^x (1 - q^{\alpha+1} - pq^{\alpha x+1})} \\ & = \frac{\{q^k (1 - q^{\alpha+1} - pq^{\alpha k+1})\}^2}{q^k (1 - q^{\alpha+1} - pq^{\alpha k+1})} \\ & = q^k (1 - q^{\alpha+1} - pq^{\alpha k+1}) \end{aligned} \quad (29)$$

□

Theorem 3. Let X be any discrete random variable with support $\{0, 1, 2, \dots\}$. For any $\alpha > 0$, X follows $EG(p, \alpha)$ iff

$$h_X(x+1) - h_X(x) = \frac{pq^{\alpha x+1} (1 - q^{\alpha+1}) (q - 2q^\alpha + q^{2\alpha+1} + 2pq^\alpha)}{(1 - q^{\alpha+1})^2 - pq^{\alpha x+1} [(1 + q^\alpha)(1 - q^{\alpha+1}) + pq^{\alpha(x+1)+1}]}, \quad (30)$$

with initial condition

$$h_X(0) = \frac{p^2(1 - q^{\alpha+1})}{1 - pq - q^{\alpha+1}}. \quad (31)$$

Proof. If part: From (30), it readily follows that

$$\begin{aligned} \sum_{x=0}^k [h_X(x+1) - h_X(x)] &= \sum_{x=0}^k \frac{pq^{\alpha x+1}(1-q^{\alpha+1})(q-2q^\alpha+q^{2\alpha+1}+2pq^\alpha)}{(1-q^{\alpha+1})^2-pq^{\alpha x+1}[(1+q^\alpha)(1-q^{\alpha+1})+pq^{\alpha(x+1)+1}]} \\ \implies h_X(k+1) - h_X(0) &= \sum_{x=0}^k \left\{ \frac{p(1-q^{\alpha(x+1)+1})(1-q^{\alpha+1})}{1-q^{\alpha+1}-pq^{\alpha(x+1)+1}} - \frac{p(1-q^{\alpha x+1})(1-q^{\alpha+1})}{1-q^{\alpha+1}-pq^{\alpha x+1}} \right\} \\ \implies h_X(k+1) - h_X(0) &= \frac{p(1-q^{\alpha(k+1)+1})(1-q^{\alpha+1})}{1-q^{\alpha+1}-pq^{\alpha(k+1)+1}} - \frac{p^2(1-q^{\alpha+1})}{1-pq-q^{\alpha+1}}. \end{aligned}$$

Hence, from (31)

$$h_X(k+1) = \frac{p(1-q^{\alpha(k+1)+1})(1-q^{\alpha+1})}{1-q^{\alpha+1}-pq^{\alpha(k+1)+1}}.$$

Only if part: Can easily be checked. □

5. Estimation

Here, we discuss four different methods of estimating the parameters of the $EG(p, \alpha)$ model. Consider a random sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$ from $X \sim EG(p, \alpha)$. A realization on \mathbf{X} is denoted by $\mathbf{x} = (x_1, x_2, \dots, x_n)$. From the viewpoint of application, \mathbf{x} is the working dataset.

5.1. Method of moment estimation

Let $\bar{x} = (1/n) \sum_{i=1}^n x_i$ be the sample mean of \mathbf{x} . Also let $s_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)$ be the sample variance. Now, from (13), $E(X) = \mu_X(q, \alpha)$, a function of q and α , is

$$\mu_X(q, \alpha) = \frac{q(1-3q^{\alpha+1}+2q^{\alpha+2}-q^{\alpha+3}+q^{2\alpha+2})}{(1-q)(1-q^{\alpha+1})(1-q+q^2-q^{\alpha+1})}.$$

Similarly from (15), $Var(X) = \sigma_X^2(q, \alpha)$ is

$$\sigma_X^2(q, \alpha) = \frac{q}{(1-q)^2} + \frac{q^{\alpha+1}}{(1-q^{\alpha+1})^2 - \frac{2(1-q)+q^2-q^{\alpha+1}}{1-q+q^2-q^{\alpha+1}}}.$$

Solving $\mu_X(q, \alpha) = \bar{x}$ and $\sigma_X^2(q, \alpha) = s_x^2$ in q and α , directly is not feasible. Now, we rewrite $\mu_X(q, \alpha)$ and $\sigma_X^2(q, \alpha)$ by transforming α to $\beta = q^\alpha$ for Algorithm 2.

$$\begin{aligned} \mu_X(q, \beta) &= \frac{q(1-3q\beta+2q^2\beta-q^3\beta+q^2\beta^2)}{(1-q)(1-q\beta)(1-q+q^2-q\beta)} \\ \sigma_X^2(q, \beta) &= \frac{q}{(1-q)^2} + \frac{q\beta}{(1-q\beta)^2 - \frac{2(1-q)+q^2-q\beta}{1-q+q^2-q\beta}} \end{aligned}$$

Algorithm 2

(For estimating (p, α) by method of moments)

- *Step-1:* Compute \bar{x} and s_x^2 from \mathbf{x} .
- *Step-2:* Fix $q \in \mathbf{q} = \{0.0001, 0.0002, \dots, 0.9999\}$.
- *Step-3:* For each $q \in \mathbf{q}$, find $\beta \in \mathbf{\beta} = \{0.0001, 0.0002, \dots, 0.9999, 1.0000\}$ such that $D(q, \beta) = (\mu_X(q, \beta) - \bar{x})^2 + (\sigma_X^2(q, \beta) - s_x^2)^2$ is minimum.
- *Step-4:* For each pair of $(q, \beta) \in (\mathbf{q}, \mathbf{\beta})$, obtained in *Step-3*, find $(\hat{q}_M, \hat{\beta}_M)$ for which $D(q, \beta)$ is minimum.
- *Step-5:* Return $\hat{p}_M = 1 - \hat{q}_M$ and $\hat{\alpha}_M = \log(\hat{\beta}_M) / \log(\hat{q}_M)$.

5.2. Estimation based on proportion of zeros and ones

Let $f(0) = \sum_{i=1}^n I(x_i = 0)$ and $f(1) = \sum_{i=1}^n I(x_i = 1)$, where $I(A) = 1$ if the statement A is true and $I(A) = 0$, else. From (6),

$$\begin{aligned} k(x) &= \frac{p_X(x+1)}{p_X(x)} = \frac{q(1-q^{\alpha(x+1)+1})}{1-q^{\alpha x+1}} \quad \text{for } x = 0, 1, \dots \\ \implies k(0) &= \frac{q(1-q^{\alpha+1})}{1-q}. \end{aligned} \quad (32)$$

From (32), we get

$$\alpha = \frac{\log \left[1 - \frac{(1-q)k(0)}{q} \right]}{\log(q)} - 1. \quad (33)$$

Algorithm 3

(For estimating (p, α) by method of relative frequency)

- *Step-1:* Compute $f(0)$ and $f(1)$ from \mathbf{x} .
- *Step-2:* Fix $q \in \mathbf{q} = \{0.0001, 0.0002, \dots, 0.9999\}$.
- *Step-3:* For each $q \in \mathbf{q}$, find α from (33), replacing $k(0)$ by $f(1)/f(0)$.
- *Step-4:* For each pair of (q, α) , obtained in *Step-3*, find $(\hat{q}_F, \hat{\alpha}_F)$ for which $(p_X(0) - f(0)/n)^2$ is minimum.
- *Step-5:* Return $\hat{p}_F = 1 - \hat{q}_F$ and $\hat{\alpha}_F$.

5.3. Maximum likelihood estimation

From (5), the log-likelihood function is

$$l(q, \alpha | \mathbf{x}) = n \log(1-q) + n \log(1-q^{\alpha+1}) - n \log(1-q+q^2-q^{\alpha+1}) + \log q \sum_{i=1}^n x_i + \sum_{i=1}^n \log(1-q^{\alpha x_i+1}).$$

By differentiating $l(q, \alpha | \mathbf{x})$ with respect to q we obtain a score function,

$$s_1(q, \alpha | \mathbf{x}) = -\frac{n}{1-q} - n \frac{q^\alpha(1+\alpha)}{1-q^{\alpha+1}} + n \frac{1-2q+q^\alpha(1+\alpha)}{1-q(1-q+q^\alpha)} + \sum_{i=1}^n \frac{x_i}{q} - \sum_{i=1}^n \frac{q^{\alpha x_i}(1+\alpha x_i)}{1-q^{\alpha x_i+1}}.$$

Similarly, by differentiating $l(q, \alpha | \mathbf{x})$ with respect to α , we obtain the other score function as

$$s_2(q, \alpha | \mathbf{x}) = -\frac{nq^{\alpha+1} \log q}{1-q^{\alpha+1}} + \frac{nq^{\alpha+1} \log q}{1-q(1-q+q^\alpha)} - \sum_{i=1}^n \frac{q^{\alpha x_i+1} x_i \log q}{1-q^{\alpha x_i+1}}.$$

It is not feasible to solve $s_1(q, \alpha | \mathbf{x}) = 0$ and $s_2(q, \alpha | \mathbf{x}) = 0$ in q and α explicitly. In absence of closed form solutions, a common practice is to numerically maximize $l(q, \alpha | \mathbf{x})$ with respect to q and α . For initialization, required for numerical optimization routines, the estimates discussed in Section 5.1 and Section 5.2 are found suitable. However, Algorithm 3 is computationally less intensive compared to Algorithm 2. If one can afford such exhaustive searching for mere initialization, it is reasonable to employ similar exhaustive searching, directly for finding maximum likelihood estimates. This will also remove the issue of instability that might creep in numerical optimization routines. As in Section 5.1, transforming α to $\beta = q^\alpha$, we obtain

$$l(q, \beta | \mathbf{x}) = n \log(1-q) + n \log(1-q\beta) - n \log(1-q+q^2-q\beta) + \log q \sum_{i=1}^n x_i + \sum_{i=1}^n \log(1-q\beta^{x_i}).$$

Algorithm 4

(For estimating (p, α) by method of maximum likelihood)

- *Step-1*: Fix $q \in \mathbf{q} = \{0.0001, 0.0002, \dots, 0.9999\}$.
- *Step-2*: For each $q \in \mathbf{q}$, find $\beta \in \mathbf{\beta} = \{0.0001, 0.0002, \dots, 0.9999, 1.0000\}$ such that $l(q, \beta | \mathbf{x})$ is maximum.
- *Step-3*: For each pair of $(q, \beta) \in (\mathbf{q}, \mathbf{\beta})$, obtained in *Step-2*, find $(\hat{q}_L, \hat{\beta}_L)$ for which $l(q, \beta | \mathbf{x})$ is maximum.
- *Step-4*: Return $\hat{p}_L = 1 - \hat{q}_L$ and $\hat{\alpha}_L = \log(\hat{\beta}_L) / \log(\hat{q}_L)$.

The second-order partial derivatives are given as follows.

$$\begin{aligned}
\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q^2} &= -\frac{n}{(1-q)^2} - \sum_{i=1}^n \frac{x_i}{q^2} - \frac{n\alpha(1+\alpha)q^{\alpha-1}}{1-q^{\alpha+1}} - \frac{nq^{2\alpha}(1+\alpha)^2}{1-q^{(\alpha+1)^2}} + \sum_{i=1}^n \frac{\alpha x_i(1+\alpha x_i)}{q(q-q^{-\alpha x_i})} \\
&\quad - \sum_{i=1}^n \frac{q^{2\alpha x_i(1+\alpha x_i)^2}}{(q^{1+\alpha x_i}-1)^2} + \frac{n(1-2q+q^\alpha(1+\alpha))^2}{(1+q(q-1-q^\alpha))^2} + \frac{n(q^\alpha \alpha(1+\alpha)-2q)}{q(1+q(q-1-q^\alpha))} \\
\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial \alpha^2} &= (\log q)^2 \left[n \frac{q^{2+\alpha}(q-1)(q-1-q^2+q^{2+2\alpha})}{(q^{1+\alpha}-1)^2(q(1-q+q^\alpha)-1)^2} - \sum_{i=1}^n \frac{x_i^2 q^{1+\alpha x_i}}{(q^{1+\alpha x_i}-1)^2} \right] \\
\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q \partial \alpha} &= \frac{n}{q-q^{-\alpha}} + \sum_{i=1}^n \frac{x_i q^{\alpha x_i}}{q^{1+\alpha x_i}-1} - \frac{n(q^{1+2\alpha}(1+\alpha)\log q)}{(q^{1+\alpha}-1)^2} + \frac{nq^\alpha(1+\alpha)\log q}{q^{1+\alpha}-1} \\
&\quad - \sum_{i=1}^n \frac{q^{1+2\alpha x_i} x_i(1+\alpha x_i)\log q}{(q^{\alpha x_i+1}-1)^2} + \sum_{i=1}^n \frac{x_i(1+\alpha x_i)\log q}{q^{\alpha x_i+1}-1} \\
&\quad + \frac{nq^{1+\alpha}\log q(1-2q+q^\alpha(1+\alpha))}{(q(1-q+q^\alpha)-1)^2} - \frac{nq^\alpha(1+(1+\alpha)\log q)}{q(1-q+q^\alpha)-1}
\end{aligned}$$

The Fisher's information matrix for (q, α) is

$$J = \begin{pmatrix} -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q^2}\right) & -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q \partial \alpha}\right) \\ -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q \partial \alpha}\right) & -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial \alpha^2}\right) \end{pmatrix}.$$

This can be approximated as

$$\hat{J} = \begin{pmatrix} J_{11} = -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q^2}\right) \Big|_{\hat{q}_L, \hat{\alpha}_L} & J_{12} = -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q \partial \alpha}\right) \Big|_{\hat{q}_L, \hat{\alpha}_L} \\ J_{21} = -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial q \partial \alpha}\right) \Big|_{\hat{q}_L, \hat{\alpha}_L} & J_{22} = -E\left(\frac{\partial^2 l(q, \alpha | \mathbf{x})}{\partial \alpha^2}\right) \Big|_{\hat{q}_L, \hat{\alpha}_L} \end{pmatrix}.$$

As $n \rightarrow \infty$, the limiting distribution of $\sqrt{n}(\hat{q}_L - q, \hat{\alpha}_L - \alpha)$ is bivariate normal with mean vector $(0, 0)$ and variance-covariance matrix

$$\hat{J}^{-1} = \frac{1}{J_{11}J_{22}-J_{12}J_{21}} \begin{pmatrix} J_{11} & -J_{12} \\ -J_{21} & J_{22} \end{pmatrix}.$$

From the relation $\hat{p}_L = 1 - \hat{q}_L$, it is evident that $\sqrt{n}(\hat{p}_L - p, \hat{\alpha}_L - \alpha)$ asymptotically follows bivariate normal distribution with mean vector $(0, 0)$ and variance-covariance matrix

$$\hat{D} = \frac{1}{J_{11}J_{22}-J_{12}J_{21}} \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}.$$

5.4. Bayesian estimation

When prior information are available, the Bayesian method of estimation with informative prior distributions on the parameters, proves to be efficient in many applications. However, in absence of the reasonable prior guess, one may assume non-informative flat priors on the parameters and compute the posteriors. Here, the two parameters are p with support $(0, 1)$ and α with support $(0, \infty)$. For the informative case, beta prior on p and gamma prior on α are reasonable. Unfortunately, with these prior choices, the posterior and hence the full conditionals though proper, are not in any identifiable family of distributions. Thus, for sampling from the marginal posteriors, the hamiltonian markov chain technique implemented in STAN (Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, Guo, Li, and Riddell 2017) is a viable alternative. Obviously, the proposed model is not available in STAN as a standard model. First, we present the function for writing the proposed model in STAN below.

```
functions{
  real EG_log(vector x, real p, real alpha){
    vector[num_elements(x)] prob;
    real lprob;
    for (i in 1:num_elements(x)){
      prob[i] <- p*(1-p)^x[i]*(1-(1-p)^(alpha*x[i]+1))*
        (1/(((1-p*(1-p))-(1-p)^(alpha+1))/(1-(1-p)^(alpha+1))));
    }
    lprob=sum(log(prob));
    return lprob;
  }
}
```

For both the informative and non-informative cases, the above code for the data-model remains same. For the informative case, we set

$$\begin{aligned} p &\sim \text{Beta}(\delta_0, \tau_0) \quad \text{and} \\ \alpha &\sim \text{Gamma}(\alpha_0, \beta_0). \end{aligned}$$

Now, we present the remaining part of the STAN code below, comprising of the ‘data’ part containing information on the hyper-parameters.

```
data{
  int<lower=0> n;
  vector[n] y;
  real<lower=0> alpha0;
  real<lower=0> beta0;
  real<lower=0> delta0;
  real<lower=0> tau0;
}
parameters{
  real<lower=0,upper=1> p;
  real<lower=0> alpha;
}
```

```

model{
  p~beta(delta0 , tau0 );
  alpha~gamma(alpha0 , beta0 );
  y~EG(p , alpha );
}

```

The information required for the data-block are to be supplied through R, interfacing STAN with R by the RStan library. For the non-informative case, we set $\pi(p, \alpha) \propto 1$, where π denotes the prior density. For this case, the remaining part of the STAN code is given below.

```

data{
  int<lower=0> n;
  vector[n] y;
}
parameters{
  real<lower=0,upper=1> p;
  real<lower=0> alpha;
}
model{
  y~EG(p , alpha );
}

```

After obtaining the posterior samples through specified number of chains, usually initial half of the samples are discarded and the estimates of p and α , namely \hat{p}_B and $\hat{\alpha}_B$ are obtained by averaging the corresponding posterior samples.

6. Applications

Here, two datasets are considered for modelling application. The first dataset records the number of claims made by automobile insurance policy holders ([Klugman, Panjer, and Willmot 2012](#)). The second one records the number of ticks counted on sheeps ([Fisher 1941](#)). For comparing the proposed model, we consider the following discrete distributions with the same support.

- $NB(r, q)$: The negative binomial distribution discussed in [Johnson, Kemp, and Kotz \(2005\)](#) having the following mass function.

$$\binom{r+x-1}{x} q^r (1-q)^x$$

- $ND(\alpha, \theta)$: The new discrete distribution discussed in [Gómez-Déniz, Sarabia, and Ojeda \(2011\)](#) having the following mass function.

$$\frac{\log(1 - \alpha\theta^x) - \log(1 - \alpha\theta^{x+1})}{\log(1 - \alpha)}$$

- $NGPL(\alpha, \theta)$: The new generalized Poisson-Lindley distribution discussed in [Bhati et al. \(2015\)](#) having the following mass function.

$$\frac{\theta^2}{(\theta + \alpha)(1 + \theta)^{x+1}} \left(1 + \frac{\alpha(x+1)}{1 + \theta} \right)$$

- $WG(\alpha, q)$: The weighted geometric distribution discussed in [Bhati and Savitri \(2018\)](#) having the following mass function.

$$(1 - \alpha)(1 - \alpha^{q+1})\alpha^x \frac{(1 - \alpha^{q(x+1)})}{1 - \alpha^q}$$

Table 1: Number of claims in automobile insurance

| Claims | Frequency | $EG(p, \alpha)$ | $NB(r, q)$ | $ND(\alpha, \theta)$ | $NGPL(\alpha, \theta)$ | $WG(\alpha, q)$ |
|----------------|-----------|-----------------|----------------|----------------------|------------------------|-----------------|
| 0 | 1563 | 1564.68 | 1564.54 | 1563.70 | 1564.57 | 1564.27 |
| 1 | 271 | 263.32 | 264.58 | 266.15 | 264.28 | 265.12 |
| 2 | 32 | 38.48 | 39.44 | 38.75 | 39.69 | 39.05 |
| 3 | 7 | 5.57 | 5.66 | 5.50 | 5.59 | 5.62 |
| 4 | 2 | 0.96 | 0.78 | 0.90 | 0.87 | 0.94 |
| 1875 | | 1875 | 1875 | 1875 | 1875 | 1875 |
| (χ^2, df) | | (2.742, 2) | (3.786, 2) | (3.002, 2) | (3.486, 2) | (2.938, 2) |
| p -value | | 0.2538 | 0.1510 | 0.2230 | 0.1760 | 0.2300 |
| MLE | | (0.145, 3.577) | (1.309, 0.871) | (-0.454, 0.141) | (8.835, 7.874) | (0.873, 0.143) |
| Variance | | (0.0002, 1.618) | (1.081, 0.214) | (0.495, 0.019) | (2.601, 0.439) | (0.683, 0.022) |
| AIC | | 1990.58 | 1991.00 | 1990.78 | 1991.18 | 1990.77 |

Table 2: Number of ticks on sheep

| Ticks | Frequency | $EG(p, \alpha)$ | $NB(r, q)$ | $ND(\alpha, \theta)$ | $NGPL(\alpha, \theta)$ | $WG(\alpha, q)$ |
|----------------|-----------|------------------|------------------|----------------------|------------------------|------------------|
| 0 | 4 | 3.17 | 5.26 | 5.46 | 7.61 | 5.36 |
| 1 | 5 | 7.89 | 7.35 | 7.12 | 7.66 | 7.72 |
| 2 | 11 | 9.21 | 8.03 | 7.75 | 7.53 | 8.41 |
| 3 | 10 | 8.99 | 7.96 | 7.76 | 7.25 | 8.21 |
| 4 | 9 | 8.15 | 7.48 | 7.40 | 6.83 | 7.57 |
| 5 | 11 | 7.12 | 6.80 | 6.81 | 6.31 | 6.75 |
| 6 | 3 | 6.10 | 6.04 | 6.12 | 5.72 | 5.90 |
| 7 | 5 | 5.16 | 5.28 | 5.40 | 5.10 | 5.08 |
| 8-10 | 7 | 11.01 | 11.77 | 12.12 | 11.69 | 11.10 |
| 11-14 | 9 | 7.88 | 8.69 | 8.94 | 8.84 | 8.16 |
| ≥ 15 | 8 | 7.32 | 7.36 | 7.16 | 7.45 | 7.74 |
| 82 | | 82 | 82 | 82 | 82 | 82 |
| (χ^2, df) | | (7.203, 8) | (9.124, 8) | (9.844, 8) | (12.666, 8) | (8.476, 8) |
| p -value | | 0.5148 | 0.3320 | 0.2761 | 0.1239 | 0.3884 |
| MLE | | (0.833, 2.786) | (1.777, 0.271) | (1.276, 0.311) | (2.312, 0.808) | (0.834, 1.759) |
| Variance | | (0.0003, 0.0422) | (0.1211, 0.0034) | (0.1600, 0.0038) | (0.6839, 0.0010) | (0.0008, 2.5250) |
| AIC | | 477.92 | 479.92 | 480.88 | 483.44 | 478.98 |

The method of maximum likelihood estimation has been employed for estimating (p, α) of the proposed distribution for modelling the datasets. The results regarding modelling the number of claims and the number of ticks are presented in Table 1 and Table 2, respectively.

As pointed out earlier in Section 2, the geometric distribution is a particular case of the proposed extended geometric distribution when $\alpha = 0$. Thus, for a given dataset, a natural question regarding the necessity of the additional parameter can be settled by testing the following hypothesis.

$$H_0 : \alpha = 0 \quad \text{versus} \quad H_1 : \alpha > 0 \quad (34)$$

For testing the hypothesis in (34), we adopt the likelihood ratio test (LRT). The corresponding test-statistic, $\Lambda = -2(l(\tilde{p}, 0) - l(\hat{p}_L, \hat{\alpha}_L)) \sim \chi^2$ distribution with degree of freedom (df) 1. Here, $\tilde{p} = 1/(1 + \bar{x})$, is the maximum likelihood estimate of p , under H_0 . H_0 is rejected at 5% level of significance when Λ exceeds 3.841. For the dataset in Table 1, $\Lambda = 1.2502$ and for the dataset in Table 2, $\Lambda = 10.434$. Thus, for the first dataset, the additional parameter has no significant role. Still, from Table 1, it is observed that, the proposed model marginally beats its closest competitors in terms of χ^2 and AIC. On the contrary, for the second dataset, the additional parameter in the proposed model, plays an important role. Thus, from Table 2, it is also seen that, the proposed model beats its closest competitors in terms of χ^2 and AIC, substantially.

Now we consider estimation of p and α for the two datasets under the Bayesian framework. We do not have any prior guess about the parameters for both the datasets. Thus, for the sake of demonstration only, we estimate p and α from Bayesian perspective using the flat prior. The flat prior considered here is $\pi(p, \alpha) \propto 1$, a non-informative prior, mentioned in Section 5.4. For a detailed elicitation on the choice of priors, one may refer to Robert (2007). For both the applications, we set the number of iterations to be 5000 and warm-up length to be 2500. The ‘iteration’ and ‘warm-up’ are two important arguments of the *stan* function in R. While the former indicates the number of samples to be generated for each one of the Markov chains, the later stands for the size of the initial sample to be discarded. Multiple Markov chains are simulated simultaneously to analyze issues related to convergence. We fix the number of chains to be 4 and note that the Rhat value in both the cases turn out to be 1. Rhat is a measure of convergence diagnostic. Rhat compares the between- and within-chain estimates for model parameters. Rhat exceeding 1 is an indication of the chains not being mixed up well. Thus, it is warranted that the posteriors are proper and the chains are convergent. For the dataset in Table 1, we obtain $\hat{p}_B = 0.187$ and $\hat{\alpha}_B = 7.411$. For the dataset in Table 2, we obtain $\hat{p}_B = 0.726$ and $\hat{\alpha}_B = 3.424$. The within chain variability of $\hat{\alpha}_B$ is found to be higher compared to \hat{p}_B .

7. Remarks

Azzalini’s skewing technique is used for the first time for generalizing a discrete distribution in this work, by considering the geometric distribution as baseline. Additionally, we have also presented a genesis of the proposed extension. The proposed model is very flexible and possesses compact forms for pmf, cdf, survival function, probability generating function and moments. It is an over dispersed distribution and can be long tailed, zero inflated, zero deflated for appropriate choices of the parameters. Three useful characterizations of the new distribution have been proved. Parameter estimation is discussed in detail and the performance is found to be satisfactory in applications. Thus, it is envisaged that the proposed two-parameter model will find its application in count data modelling. Of course, the idea can be used for other count distributions as well. Further work on numerical comparison of performances of different estimation procedures is warranted. Modelling count data with covariates using appropriate re-parameterized version of the proposed model will also be taken

up in a follow-up work.

Acknowledgement

The authors would like to thank the editor and the anonymous reviewers for their insightful comments and suggestion, which have improved the paper. S. Chakraborty and S. H. Ong are supported by the Ministry of Higher Education Grant FRGS/1/2020/STG06/SYUC/02/1.

References

- Azzalini A (1985). “A Class of Distributions which Includes the Normal Ones.” *Scandinavian Journal of Statistics*, **12**(2), 171–178.
- Bhati D, Sastry DVS, Maha Qadri PZ (2015). “A New Generalized Poisson Lindley Distribution.” *Austrian Journal of Statistics*, **44**(4), 35–51. doi:10.17713/ajs.v44i4.54.
- Bhati D, Savitri J (2018). “Weighted Geometric Distribution with New Characterizations of Geometric Distribution.” *Communications in Statistics - Theory and Methods*, **47**(6), 1510–1527. doi:10.1080/03610926.2017.1321765.
- Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A (2017). “Stan: A Probabilistic Programming Language.” *Journal of Statistical Software*, **76**(1), 218–223. doi:10.18637/jss.v076.i01.
- Chakraborty S, Gupta RD (2015). “Exponentiated Geometric Distribution: Another Generalization of Geometric Distribution.” *Communications in Statistics – Theory and Methods*, **44**(6), 1143–1157. doi:10.1080/03610926.2012.763090.
- Fisher RA (1941). “The Negative Binomial Distribution.” *Annals of Eugenics*, **11**(1), 182–187.
- Gómez-Déniz E (2010). “Another Generalization of the Geometric Distribution.” *Test*, **19**(2), 399–415. doi:10.1007/s11749-009-0169-3.
- Gómez-Déniz E, Sarabia JM, Ojeda EC (2011). “A New Discrete Distribution with Actuarial Applications.” *Insurance: Mathematics and Economics*, **48**(3), 406–412. doi:10.1016/j.insmatheco.2011.01.007.
- Jain GC, Consul PC (1971). “A Generalized Negative Binomial Distribution.” *SIAM Journal on Applied Mathematics*, **21**(4), 501–513. doi:10.1137/0121056.
- Johnson NL, Kemp AW, Kotz S (2005). *Univariate Discrete Distributions*. Wiley. ISBN 978-0-471-27246-5.
- Keilson J, Gerber H (1971). “Some Results for Discrete Unimodality.” *Journal of American Statistical Association*, **66**(334), 386–389. doi:10.1080/01621459.1971.10482273.
- Klugman SA, Panjer HH, Willmot GE (2012). *Loss Models: From Data to Decisions*. John Wiley & Sons. ISBN 978-1-119-52378-9.
- Makcutek J (2008). “A Generalization of the Geometric Distribution and Its Application in Quantitative Linguistics.” *Romanian Reports in Physics*, **60**(3), 501–509.
- Nekoukhou V, Alamatsaz MH, Bidram H (2012). “A Discrete Analogue of the Generalized Exponential Distribution.” *Communications in Statistics - Theory and Methods*, **41**(11), 2000–2013. doi:10.1080/03610926.2011.555044.

- Ong SH, Muthaloo S (2012). "A Class of Discrete Distributions Suited to Fitting Very Long-Tailed Data." *Communications in Statistics - Simulation and Computation*, **41**(4), 929–945. doi:10.1080/03610919508813285.
- Philippou AN, Georgiou C, Philippou GN (1983). "A Generalized Geometric Distribution and Some of Its Properties." *Statistics and Probability Letters*, **1**(4), 171–175. doi:10.1016/0167-7152(83)90025-1.
- Robert CP (2007). *The Bayesian Choice: From Decision-theoretic Foundations to Computational Implementation*. Springer. ISBN 978-0-387-71599-5.
- Tripathi RC, Gupta RC, White TJ (1987). "Some Generalizations of the Geometric Distribution." *Sankhya Ser. B*, **49**(3), 218–223.

Affiliation:

Aniket Biswas
 Department of Statistics
 Dibrugarh University
 Dibrugarh, Assam, India- 786004. E-mail: biswasaniket44@gmail.com
 URL: <https://dibru.ac.in/userlist/welcome/profile/569>

Subrata Chakraborty
 Department of Statistics
 Dibrugarh University
 Dibrugarh, Assam, India- 786004. E-mail: subrata_stats@dibru.ac.in
 URL: <https://dibru.ac.in/userlist/welcome/profile/239>

Seng Huat Ong
 Department of Actuarial Science and Applied Statistics
 UCSI University
 Kuala Lumpur, Malaysia-56000. E-mail: ongsh@ucsiuniversity.edu.my
 URL: <https://www.ucsiuniversity.edu.my/professor-dr-ong-seng-huat-fasc>