İSTATİSTİK: JOURNAL OF THE TURKISH STATISTICAL ASSOCIATION
Vol. 14, No. 1, July 2022, pp. 1–10
ISSN 1300-4077 | 22 | 1 | 1 | 10

İSTATİSTİK

HYPOGEOMETRIC DISTRIBUTION AND RELATED DISCRETE TIME POINT PROCESS

Stefanka Chukova*
School of Mathematics and Statistics,
Victoria University of Wellington,
Wellington, NZ

Leda Minkova and Silvana Paralloi
Faculty of Mathematics and Informatics,
Sofia University,
Sofia, Bulgaria

Abstract: In this paper we propose and study a new distribution, called the hypogeometric distribution, which is a sum of independent geometrically distributed variables with different parameters. Also, we propose and study a discrete time point process based on this distribution. As an example, we focus on a particular form of this process. Also, we show that this type of processes could be used as an appropriate tool to model arrivals with increasing or decreasing time trends. Some possible extensions of this work are also included in the paper.

Key words: Geometric distribution, Hypogeometric distribution, Waiting time, Counting process.

1. Introduction

Most theoretical risk models are formulated for continuous time and results of interest for the particular study are derived. On the other hand, the practical world is discrete, and the continuous time models have to be modified and adjusted to the discrete time scenario. The results for discrete time risk models can provide a good background for better understanding the ideas of the continuous-time scenario and their results can be used as approximations or bounds for the corresponding results in the continuous case, see [3] and [2] for the approximating procedures. The discrete-time risk models have their special features and require specific set of ideas and apparatus to analyze. Also, they are of independent interest since formulas for discrete-time models are mostly recursive and hence suitable for computing the quantities of interest in practice while still reproducing, in limit, the corresponding continuous time results.

It is well-known that if the counting process in the discrete time risk model is the binomial process, the interarrival times are independent, identically distributed geometric random variables, see for example [5]. In this paper we consider a point process, with interarrival times that are independent, geometrically distributed with different parameters random variables. The geometric summands with different parameters are used in [1] and [6] for representing the number of shocks in an engineering system.

The main goal of this paper is to introduce the discrete-time hypogeometric process (HPGP), which has similar structure as the Binomial process, but the interarrival times are not identically distributed. In order to define this process, we firstly introduce the hypogeometric distribution, which is an analogue to the hypoexponential distribution given in [7], and use it to propose and study the HPGP.

In the next Section 2, we introduce the hypogeometric distribution. The corresponding discrete time pure birth process with some properties is introduced in Section 3. In Section 4 we illustrate

* Corresponding author. E-mail address:leda@fmi.uni-sofia.bg

our ideas on an example. The discussion in Section 5 provides some pictorial illustrations of the defined process and discuss some of its properties. Some conluding remarks for this study are given in Section 6.

2. Hypogeometric distribution

Let us consider the sequence

$$\{X_j,\ j=1,2,\ldots\} \tag{2.1}$$

of mutually independent, geometrically distributed random variables with different parameters $\pi_j \in (0,1)$ with $\pi_i \neq \pi_j,\ i \neq j$. The corresponding probability mass function (PMF), cumulative distribution function (CDF) and probability generating function (PGF) of $X_j,\ j=1,2,\ldots$ in (2.1) are given by

$$P(X_j=n)=(1-\pi_j)\pi_j^{n-1},\ n=1,2,\ldots, \tag{2.2}$$

$$F_{X_j}(n)=P(X_j\leq n)=1-\pi_j^n,\ n=1,2,\ldots, \tag{2.3}$$

and

$$\Psi_{X_j}(s)=\frac{(1-\pi_j)s}{1-\pi_j s},\ j=1,2,\ldots$$

Let $\tau_k=X_1+\ldots+X_k,\ k=1,2,\ldots$ be the sum of k of these random variables. Then, the following lemma holds:

Lemma 1. *The PMF of $\tau_k,\ k=1,2,\ldots$ is given by*

$$P(\tau_k=n)=\begin{cases}0, & n<k\\ \sum_{j=1}^{k} w(k,j)P(X_j=n), & n=k,k+1,\ldots,\end{cases} \tag{2.4}$$

where

$$w(k,j)=\Pi_{i=1,i\neq j}^{k}\frac{1-\pi_i}{\pi_j-\pi_i},\ j=1,2,\ldots,k, \tag{2.5}$$

and $w(1,1)=1$.

Proof. According to the definition, the PGF of τ_k is given by

$$\Psi_{\tau_k}(s)=s^k\ \Pi_{j=1}^{k}\frac{1-\pi_j}{1-\pi_j s}=s^k\,\Phi(s),$$

where

$$\Phi(s)=\Pi_{j=1}^{k}\frac{1-\pi_j}{1-\pi_j s}=\frac{\Pi_{j=1}^{k}(\frac{1}{\pi_j}-1)}{\Pi_{j=1}^{k}(\frac{1}{\pi_j}-s)}=\frac{\tilde{H}(s)}{D(s)}$$

is the PGF of some random variable Y with distribution $P(Y=n)=p_n,\ n=0,1,\ldots$. Then, due to the properties of PGF, we have the following representation for τ_k

$$\tau_k=Y+k,$$

i.e., $P(\tau_k=n)=P(Y=n-k)=p_{n-k},\ n\geq k$. Therefore, to find the distribution of τ_k it suffices to find the distribution of Y. We invert $\Phi(s)$ by using the partial-fraction expansion method, given in [4], p.220.

The roots of the denominator $D(s)=\Pi_{j=1}^{k}(\frac{1}{\pi_j}-s)$ are $s_j=\frac{1}{\pi_j},\ j=1,2,\ldots,k$ and its derivative is given by $D'(s)=-\sum_{i=1}^{k}\Pi_{j=1,j\neq i}^{k}(\frac{1}{\pi_j}-s)$. Then

$$D'(s_i)=D'\left(\frac{1}{\pi_i}\right)=-\Pi_{j=1,j\neq i}^{k}\left(\frac{1}{\pi_j}-\frac{1}{\pi_i}\right).$$

Then, for the coefficients f_i, we have

$$f_i = -\frac{\tilde{H}(s_i)}{D'(s_i)} = \frac{\Pi_{j=1}^{k}(\frac{1}{\pi_j} - 1)}{\Pi_{j=1, j\neq i}^{k}(\frac{1}{\pi_j} - \frac{1}{\pi_i})},\ i = 1, 2, \ldots k.$$

Hence, according to the inversion method, the PMF of the random variable Y is as follows

$$\begin{aligned} P(Y = n) &= \sum_{i=1}^{k} \frac{\Pi_{j=1}^{k}(\frac{1}{\pi_j} - 1)}{\Pi_{j=1, j\neq i}^{k}(\frac{1}{\pi_j} - \frac{1}{\pi_i})} \pi_i^{n+1} \\ &= \sum_{i=1}^{k} \frac{\Pi_{j=1}^{k}(1 - \pi_j)}{\Pi_{j=1, j\neq i}^{k}(1 - \frac{\pi_j}{\pi_i})} \pi_i^{n} \\ &= \sum_{i=1}^{k} \Pi_{j=1, j\neq i}^{k} \left(\frac{1 - \pi_j}{\pi_i - \pi_j}\right)(1 - \pi_i)\pi_i^{n+k-1} \\ &= \sum_{i=1}^{k} w(k, i) P(X_i = n + k),\ n = 0, 1, \ldots \end{aligned} \tag{2.6}$$

Now, using the distribution of Y in (2.6), we find that the PMF of τ_k is given by

$$P(\tau_k = n) = P(Y = n - k) = \begin{cases} 0, & n < k \\ \sum_{j=1}^{k} w(k, j) P(X_j = n), & n = k, k+1, \ldots \end{cases}$$

□

Remark 1. Due to (2.4), the following identity is true

$$P(\tau_{k+1} = n) = \sum_{j=1}^{k+1} w(k+1, j) P(X_j = n) = 0,\ n = 1, 2, \ldots, k. \tag{2.7}$$

Remark 2. The following identities are true

$$w(k, j) = \frac{\pi_j - \pi_{k+1}}{1 - \pi_{k+1}} w(k+1, j),\ j = 1, 2, \ldots k+1. \tag{2.8}$$

and

$$\sum_{j=1}^{k} w(k, j) \pi_j^{k-1} = 1,\ k = 1, 2, \ldots \tag{2.9}$$

The equation (2.9) is equivalent to the fact that $\sum_{n=k}^{\infty} P(\tau_k = n) = 1,\ k = 1, 2, \ldots$

Definition 1. The distribution of τ_k, given in (2.4) is called a hypogeometric distribution with parameters $\pi_1, \pi_2, \ldots, \pi_k$, $\pi_1 \neq \pi_2 \neq \ldots \neq \pi_k$, and it is denoted by $HPG(\pi_1, \pi_2, \ldots, \pi_k)$.

Lemma 2. *The cumulative distribution function of* $\tau_k,\ k = 1, 2, \ldots$ *is given by*

$$P(\tau_k \leq n) = \sum_{j=1}^{k} w(k, j) \pi_j^{k-1}(1 - \pi_j^{n-k+1}),\ k \leq n. \tag{2.10}$$

PROOF. According to the definition, we have the following

$$P(\tau_k \le n) = \sum_{i=k}^{n}\sum_{j=1}^{k} w(k,j)P(X_j = i)$$

$$= \sum_{j=1}^{k} w(k,j)\sum_{i=k}^{n} P(X_j = i)$$

$$= \sum_{j=1}^{k} w(k,j)(1-\pi_j)\sum_{i=k}^{n} \pi_j^{i-1},$$

which leads to (2.10). □

3. Point process with $HPG(\pi_1, \pi_2, \ldots, \pi_k)$ distributed k^{th} waiting time

Let us consider a point process, whose waiting times are $HPG(\pi_1, \pi_2, \ldots, \pi_k)$. For this process, we will use the following notation $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$. Then, it is well-known that the expected value and the variance of the j^{th} interevent time are equal to

$$E(X_j) = \frac{1}{1-\pi_j} \quad \text{and} \quad V(X_j) = \frac{\pi_j}{(1-\pi_j)^2}. \tag{3.1}$$

For this process, let us denote by $N(n)$ the number of events up to and including time n, with $N(0) = 0$. Then the following theorem holds:

THEOREM 1. *The PMF of $N(n)$, $n = 1, 2, \ldots$ is given by*

$$P(N(n) = k) = \begin{cases} \pi_1^n, & k = 0, \\ \frac{1}{1-\pi_{k+1}} \sum_{j=1}^{k+1} w(k+1,j)P(X_j = n+1), & k = 1, 2, \ldots, n. \end{cases} \tag{3.2}$$

PROOF. Firstly, let $k = 0$. The events $\{N(n) = 0\} \equiv \{X_1 > n\}$ are equivalent, i.e., $P(N(n) = 0) = P(X_1 > n)$. But $P(X_1 > n) = \pi_1^n$, therefore (3.2) is true for $k = 0$.

Let $k = 1, 2, \ldots, n$. According to the well-known relation

$$P(N(n) = k) = P(\tau_k \le n) - P(\tau_{k+1} \le n),$$

we have that

$$P(N(n) = k) = \sum_{j=1}^{k} w(k,j)\pi_j^{k-1}(1-\pi_j^{n-k+1}) - \sum_{j=1}^{k+1} w(k+1,j)\pi_j^{k}(1-\pi_j^{n-k}).$$

Then, using the identity (2.8), we have

$$P(N(n) = k) = \sum_{j=1}^{k} w(k+1,j)\frac{\pi_j - \pi_{k+1}}{1-\pi_{k+1}}(\pi_j^{k-1} - \pi_j^n) - \sum_{j=1}^{k} w(k+1,j)\pi_j^k(1-\pi_j^{n-k})$$

$$-w(k+1,k+1)\pi_{k+1}^k(1-\pi_{k+1}^{n-k})$$

$$= \sum_{j=1}^{k} w(k+1,j)\frac{1-\pi_j}{1-\pi_{k+1}}(\pi_j^n - \pi_{k+1}\pi_j^{k-1}) - w(k+1,k+1)(\pi_{k+1}^k - \pi_{k+1}^n)$$

$$= \sum_{j=1}^{k+1} w(k+1,j)\frac{1-\pi_j}{1-\pi_{k+1}}(\pi_j^n - \pi_{k+1}\pi_j^{k-1}).$$

According to the identity (2.7), the second part of this expression is zero, which leads to (3.2). □

Remark 3. The PMF of $N(n)$ for $k=1,2,\ldots,n$ given in (3.2), has the following equivalent representation:

$$P(N(n)=k)=\Pi_{m=1}^{k}(1-\pi_m)\sum_{j=1}^{k+1}\frac{\pi_j^n}{\Pi_{m=1,m\neq j}^{k+1}(\pi_j-\pi_m)}. \tag{3.3}$$

It follows from the definition of $w(k,j)$ in (2.5).

Remark 4. Due to (3.2), the mean of $N(n)$ is given by

$$E(N(n))=\sum_{k=1}^{n}\frac{k}{1-\pi_{k+1}}\sum_{j=1}^{k+1}w(k+1,j)P(X_j=n+1),\ n=1,2,\ldots$$

Next, let us assume that the state transition probabilities of a counting process $N^*(n)$, with $N^*(0)=0$, are governed by the following assumptions:

$$P(N^*(1)=k)=\begin{cases}\pi_1, & k=0,\\ 1-\pi_1, & k=1,\\ 0, & k\geq 2,\end{cases} \tag{3.4}$$

and for every $k=0,1,\ldots,$ and $n=1,2,\ldots$

$$P(N^*(n+1)=k+j\mid N^*(n)=k)=\begin{cases}\pi_{k+1}, & j=0,\\ 1-\pi_{k+1}, & j=1,\\ 0, & j\geq 2,\end{cases} \tag{3.5}$$

which defines $N^*(n)$ as a discrete pure birth process. Next, we show that the following theorem holds.

Theorem 2. *The process $N^*(n)$ defined by the assumptions (3.4) and (3.5) coincides with the counting process $N(n)$ whose interevent times are given by the sequence (2.1) with (2.2).*

Proof. It suffices to show that the distribution of $N^*(n)$ coincides with the distribution of $N(n)$ given in Theorem 1. We use mathematical induction to prove this coincidence.

For $n=0$, we have $P(N(0)=0)=P(N^*(0)=0)=1$ by definition. For $n=1$, the distribution of $N^*(1)$ is given by (3.4). The distribution of $N(1)$, using (3.2) for $n=1$ and $\{k=0,1\}$, we obtain (3.4).

For $n=2$, the distribution of $N^*(2)$, using the probability rules and the total probability rule we get

- $k=0$

$$P(N^*(2)=0)=P(N^*(2)=0|P(N^*(1)=0)P(N^*(1)=0)=\pi_1\pi_1=\pi_1^2.$$

- $k=1$

$$\begin{aligned}P(N^*(2)=1)=\ &P(N^*(2)=1|P(N^*(1)=0)P(N^*(1)=0)\\ &+P(N^*(2)=1|P(N^*(1)=1)P(N^*(1)=1)\\ &=\pi_1(1-\pi_1)+(1-\pi_1)\pi_2=(1-\pi_1)(\pi_1+\pi_2).\end{aligned}$$

- $k=2$

$$P(N^*(2)=2)=P(N^*(2)=2|P(N^*(1)=1)P(N^*(1)=1)=(1-\pi_1)(1-\pi_2).$$

For the distribution of $N(2)$, using (3.2) for $n=2$ and $\{k=0,1,2\}$, we obtain the same distribution as for $N^*(2)$.

Now, assume that the two distributions coincide for some $n=i>2$, i.e., $P(N(i)=k)=P(N^*(i)=k)$ for $k=0,1,2,\ldots,i$. Next, we show that they coincide for $n=i+1$. The distribution of $N^*(i+1)$ is as follows

- $k=0$

$$P(N^*(i+1)=0)=P(N^*(i+1)=0|P(N^*(i)=0)P(N^*(i)=0)=\pi_1\pi_1^i=\pi_1^{i+1}.$$

- For $k=1,2,\ldots,i+1$, we have

$$\begin{aligned}
P(N^*(i+1)=k) &= P(N^*(i+1)=k|P(N^*(i)=k-1)P(N^*(i)=k-1)\\
&+P(N^*(i+1)=k|P(N^*(i)=k)P(N^*(i)=k)\\
&=\Pi_{m=1}^{k-1}(1-\pi_m)\sum_{j=1}^{k}\frac{\pi_j^i}{\Pi_{m=1,m\neq j}^{k}(\pi_j-\pi_m)}(1-\pi_k)\\
&+\Pi_{m=1}^{k}(1-\pi_m)\sum_{j=1}^{k+1}\frac{\pi_j^i}{\Pi_{m=1,m\neq j}^{k+1}(\pi_j-\pi_m)}\pi_{k+1}\\
&=\Pi_{m=1}^{k}(1-\pi_m)\sum_{j=1}^{k}\frac{\pi_j^i}{\Pi_{m=1,m\neq j}^{k}(\pi_j-\pi_m)}\left(\frac{\pi_{k+1}}{\pi_j-\pi_{k+1}}+1\right)\\
&+\frac{\Pi_{m=1}^{k}(1-\pi_m)}{\Pi_{m=1}^{k}(\pi_{k+1}-\pi_m)}\pi_{k+1}^{i+1}\\
&=\Pi_{m=1}^{k}(1-\pi_m)\sum_{j=1}^{k+1}\frac{\pi_j^{i+1}}{\Pi_{m=1,m\neq j}^{k+1}(\pi_j-\pi_m)}.
\end{aligned}$$

Using (3.3) for $n=i+1$ and $\{k=1,2,\ldots,i+1\}$, we obtain that the distribution of $N(i+1)$ is the same as the distribution of $N^*(i+1)$.

□

THEOREM 3. *The counting process $N(n)$ satisfies the following recursion formula*

$$P(N(n)=k)=\begin{cases}\pi_1^n, & k=0,\\ \pi_{k+1}P(N(n-1)=k)+(1-\pi_k)P(N(n-1)=k-1), & k=1,2,\ldots,n-1\\ (1-\pi_1)\ldots(1-\pi_n), & k=n.\end{cases}$$

PROOF. The recursion follows from the assumptions (3.4) and (3.5). □

4. An example: $HPGP(\pi_1,\pi_1^a,\ldots,\pi_1^{a^{k-1}})$

According to the definition of the $HPGP(\pi_1,\pi_1^a,\ldots,\pi_1^{a^{k-1}})$, the PMF of $X_j,\ j=1,2,3,\ldots$ is given by

$$P(X_j=n)=(1-\pi_1^{a^{j-1}})\pi_1^{a^{j-1}(n-1)},\ n=1,2,\ldots \tag{4.1}$$

The mean and the variance of X_j, are given by

$$E(X_j)=\frac{1}{1-\pi_1^{a^{j-1}}}\quad\text{and}\quad V(X_j)=\frac{\pi_1^{a^{j-1}}}{(1-\pi_1^{a^{j-1}})^2}. \tag{4.2}$$

In this case, there are only two parameters that affect the behaviour of the process, that is why we denote it by $HPGP(a,\pi_1)$.

If $a>1$, it is easy to verify that $\pi_1^{a^{j-1}}\to 0$ and then $E(X_j)\to 1$ and $V(X_j)\to 0$, as $j\to\infty$. If $a<1$, $\pi_1^{a^{j-1}}\to 1$ and $E(X_j)\to\infty$ and $V(X_j)\to\infty$. It is always $a\neq 1$, due to the assumption $\pi_i\neq\pi_j,\ i\neq j$ in the initial settings of the process. If $a=1$, the random variables $X_1,X_2,\ldots$ are i.i.d., geometrically distributed as X_1, therefore $HPGP(a,\pi_1)$ is a discrete time renewal process and the corresponding counting process is the binomial process.

4.1. The k^{th} waiting time for $HPGP(a, \pi_1)$

Let τ_k is the waiting time until the occurrence of the k^{th} event in a HPGP(a, π_1), i.e. $\tau_k = X_1 + \ldots + X_k$. Therefore, using Lemma 1, we obtain the following theorem:

THEOREM 4. *The distribution of the waiting time τ_k until the k^{th} event for a $HPGP(a, \pi_1)$ is given by*

$$P(\tau_k = n) = \begin{cases} 0, & n < k \\ \sum_{j=1}^{k} w(k,j) P(X_j = n), & n = k, k+1, \ldots, \end{cases} \tag{4.3}$$

where

$$w(k,j) = \Pi_{l=1, l \neq j}^{k} \frac{1 - \pi_1^{a^{l-1}}}{\pi_1^{a^{j-1}} - \pi_1^{a^{l-1}}}, \; j = 1, 2, \ldots, k.$$

PROOF. It follows from Lemma 1 with $\pi_j = \pi_1^{a^{j-1}}$. □

Also, we get that

$$E(\tau_k) = \sum_{j=1}^{k} \frac{1}{1 - \pi_1^{a^{j-1}}} \quad \text{and} \quad V(\tau_k) = \sum_{j=1}^{k} \frac{\pi_1^{a^{j-1}}}{(1 - \pi_1^{a^{j-1}})^2}.$$

4.2. The $HPGP(a, \pi_1)$ counting process

Denote by $N(n)$ the counting process, representing the number of events in a HPGP(a, π_1) up to and including time $n \geq 0$, i.e., $N(n) = \max\{k, \tau_k \leq n\}$. The state space of $N(n)$ is $\mathcal{N}$, the set of the non-negative integers. Then we have the following result:

THEOREM 5. *The probability mass function of $N(n)$ in given by*

$$P(N(n) = k) = \begin{cases} \pi_1^n, & k = 0, \\ \frac{1}{1 - \pi_1^{a^k}} \sum_{j=1}^{k+1} w(k+1, j) P(X_j = n+1), & k = 1, 2, \ldots, n. \end{cases} \tag{4.4}$$

PROOF. The proof follows directly from Theorem 1 with $\pi_j = \pi_1^{a^{j-1}}$. □

According to Theorem 2, the assumption (3.5) is given by

$$P(N(n+1) = k + j \mid N(n) = k) = \begin{cases} \pi_1^{a^k}, & j = 0, \\ 1 - \pi_1^{a^k}, & j = 1, \\ 0, & j \geq 2, \end{cases} \tag{4.5}$$

for every $k = 0, 1, \ldots,$ and $n = 1, 2, \ldots,$ with the initial distribution given by

$$P(N(1) = k) = \begin{cases} \pi_1, & k = 0, \\ 1 - \pi_1, & k = 1, \\ 0, & k \geq 2, \end{cases} \tag{4.6}$$

which leads to the following equivalent definition of the HPGP:

DEFINITION 2. The counting process $N(n)$, $n = 1, 2, \ldots,$ defined by the assumptions (4.5) and (4.6) is a $HPGP(a, \pi_1)$.

4.3. The expected waiting time

Let us define

$$\tau_{N(n)} = \begin{cases} 0, & N(n)=0, \\ X_1 + X_2 + \ldots + X_{N(n)}, & N(n)=1,2,\ldots,n. \end{cases}$$

THEOREM 6. *The expected waiting time of $\tau_{N(n)}$ is given by*

$$E(\tau_{N(n)}) = \sum_{j=1}^{n} E(X_j) P(N(n) \geq j). \tag{4.7}$$

PROOF. For the mean of $\tau_{N(n)}$ we have

$$\begin{aligned} E(\tau_{N(n)}) &= E[E(\textstyle\sum_{i=1}^{N(n)} X_i | N(n))] \\ &= \textstyle\sum_{k=1}^{n} E(\tau_k) P(N(n)=k) \\ &= \textstyle\sum_{k=1}^{n} \sum_{j=1}^{k} \frac{1}{1-\pi_1^{a^{j-1}}} P(N(n)=k) \\ &= \textstyle\sum_{j=1}^{n} \frac{1}{1-\pi_1^{a^{j-1}}} \sum_{k=j}^{n} P(N(n)=k), \end{aligned} \tag{4.8}$$

and then (4.7).

5. Discussion on $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$

In what follows, we provide some insight on the behaviour of $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$ and its particular version $HPGP(a, \pi_1)$. Recall that the consecutive interarrival times X_j, $j = 1, 2, 3\ldots$ of the $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$ are geometrically distributed with parameter $(1 - \pi_j)$. Also, see (3.1), we have

$$E(X_j) = \frac{1}{1-\pi_j} \quad \text{and} \quad V(X_j) = \frac{\pi_j}{(1-\pi_j)^2}.$$

Let us consider the following sequence of the consecutive parameters of $HPG(\pi_1, \pi_2, \ldots, \pi_k)$

$$\begin{gathered} \pi_1 < \pi_2 < \ldots < \pi_k, \text{ k} = 2, 3, 4, \ldots, \text{ i.e.,} \\ 1-\pi_1 > 1-\pi_2 > \ldots > 1-\pi_k, \text{ k} = 2, 3, 4, \ldots \end{gathered} \tag{5.1}$$

Then, using formula (2.3) and the definition of usual stochastic order, denoted by "$\prec_{st}$", it is easy to see that

$$P(X_j > n) = \pi_j^n \leq \pi_{j+1}^n = P(X_{j+1} > n), \text{ then } X_j \prec_{st} X_{j+1}.$$

For details on stochastic orderings see [8]. Then, the consecutive interarrival times $\{X_j\}_1^\infty$ of the $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$ form a stochastically increasing sequence. Also, it is easy to see that $E(X_1) < E(X_2) < \ldots < E(X_k) < \ldots$. Therefore, $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$ can be used as a tool to model increasing trends over time.

Similarly, if

$$\begin{gathered} \pi_1 > \pi_2 > \ldots > \pi_k, \text{ k} = 2, 3, 4, \ldots, \text{ i.e.,} \\ 1-\pi_1 < 1-\pi_2 < \ldots < 1-\pi_k, \text{ k} = 2, 3, 4, \ldots, \end{gathered} \tag{5.2}$$

then the consecutive interarrival times of the process form a stochastically decreasing sequence and $E(X_1) > E(X_2) > \ldots > E(X_k) > \ldots$.

Analogously, if $a < 1$,

$$\pi_1 < \pi_1^a < \pi_1^{a^2} < \ldots$$

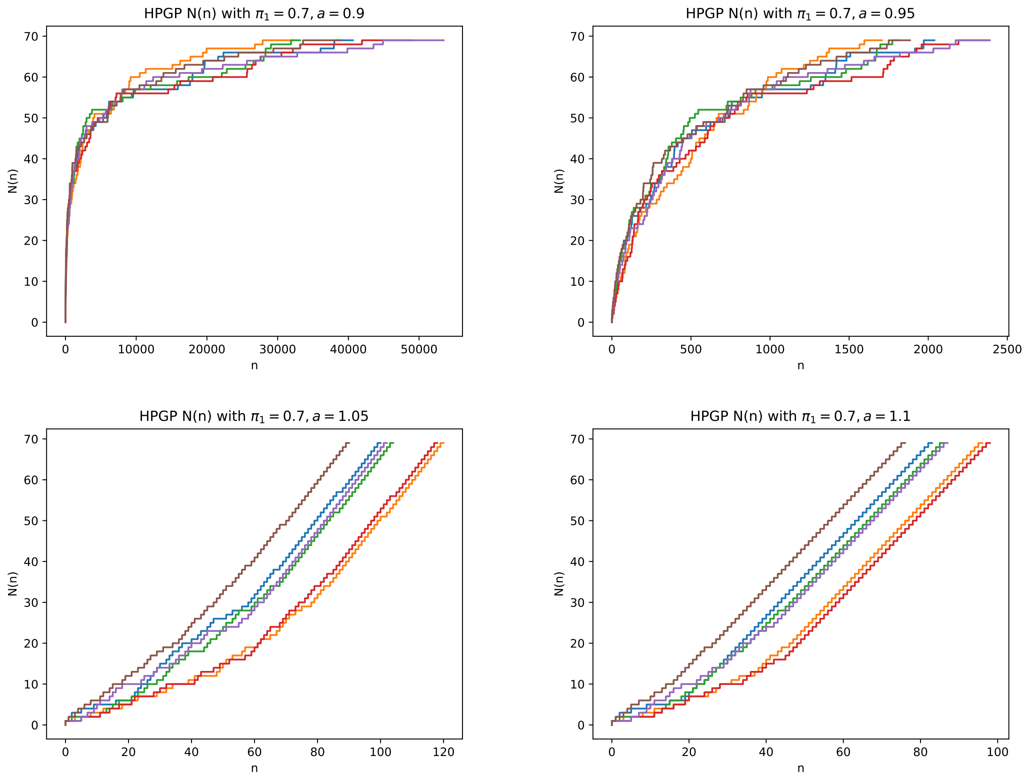

FIGURE 1. $N(n)$ for $HPGP(a, \pi_1)$ with different parameter values

and, using (5.1), we conclude that the $HPGP(a, \pi_1)$ is a stochastically increasing, whereas for $a > 1$ it is stochastically decreasing process, which is illustrated in Figure 1. In each of the plots included in Figure 1, we have shown six sample paths of the counting process of the $HPGP$ with the parameters as given in the plots' label. Comparing the two plots on the first column of Figure 1, we see that it takes much longer (it takes many more discrete time steps) for $HPGP(0.9, 0.7)$ to reach level 70 than for $HPGP(1.05, 0.7)$ to reach the same level. Similar comparison is in place for the second column plots of Figure 1.

Next, we provide some insight on the behaviour of $E(N(n))$ of $HPGP(a, \pi_1)$ depending on the values of its parameters. We use Remark 4 (and also simulation) to compute $E(N(n))$. The plots on Figure 2 agree with our intuition regarding the behaviour of $E(N(n))$. Indeed, $HPGP(0.8, 0.7)$ is a stochastically increasing process, therefore its expected number of events at time 30 should be less than corresponding number of events of $HPGP(1.10, 0.7)$, which forms a decreasing process, at the same time. So, Figure 2 depicts the corresponding $E(N(n))$'s for $HPGP$, with parameters given in the legend, having a relationship as we have expected.

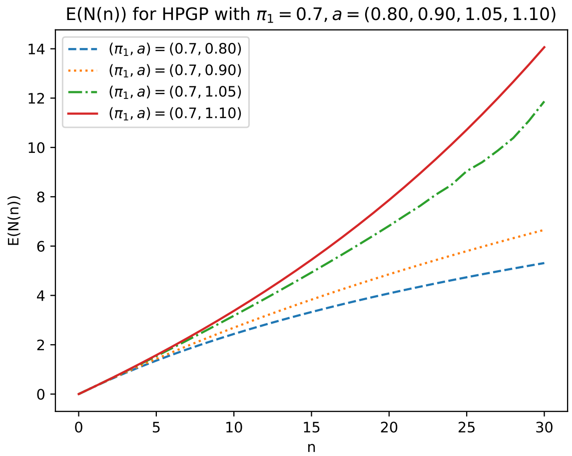

FIGURE 2. $E(N(n))$ for $HPGP(a, \pi_1)$ with different parameter values

6. Conclusion

In this study, analogueslly to the ideas of the hypoexponential distribution in [7], we proposed a new discrete distribution, called hypogeometric distribution, which is a sum of independent geometrically distributed random variables with different parameters. Also, we studied a point process with hypogeometrically distributed waiting times and derived some of its basic properties. An example of this type of process, with a particular hypogeometric distribution for its waiting times, is also included in the paper. In addition, a discussion on some useful properties of these type of processes to model time trends is included.

There are many open research questions related to the newly introduced hypogeometric distribution and related discrete-time point process. For example, questions related to the statistical inference for the distribution/process parameters as well as fitting these to real datasets. Also, how to introduce a compound $HPGP(\pi_1, \pi_2, \ldots, \pi_k)$ and $HPGP(a, \pi_1)$? What are the possible applications of these processes in risk theory?

Acknowledgment The research of the second author was partially supported by Grant DN12/11/20.dec.2017 of the Ministry of Education and Science of Bulgaria.

References

[1] 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, 161-171.

[2] Cossette, H., Landriault, D. and Marceau, É. (2004). Exact expressions and upper bound for ruin probabilities in the compound Markov binomial model. *Insurance: Mathematics and Economics*, 34, 449-466.

[3] Dickson, D.C.M., Egidio dos Reis, A. and Waters, H.R. (1995). Some stable algorithms in ruin theory and their applications. *ASTIN Bulletin*, 25, 153-175.

[4] Kobayashi, H., Mark, B.I. and Turin, W. (2012). *Probability, Random Processes, and Statistical Analysis*. Cambridge University Press, New York.

[5] Li, Sh. and Sendova, K.P. (2013). The finite-time ruin probability under the compound binomial risk model. *European Actuarial Journal*, 3, 249-271.

[6] Ortega, E-M., Alonso, J. and Ortega-Quiles, E. (2016). Comparison of multistate models with discrete-time pure-birth process for recurrent events and uncertain parameters. *Communications in Statistics - Theory and Methods*, 45, 1733-1746.

[7] Ross, S.M. (2007). *Introduction to Probability Models*. Ninth edition, Elsevier Inc., USA.

[8] Shaked, M. and Shanthikumar, J.G. (1994). *Stochastic Orders and Their Applications*. Academic Press, New York.