Methodol Comput Appl Probab (2017) 19:377–393
DOI 10.1007/s11009-016-9482-y

Compound Geometric Distribution of Order *k*

Markos V. Koutras[1] **· Serkan Eryilmaz**[2]

Received: 11 February 2015 / Revised: 7 January 2016 /
Accepted: 15 January 2016 / Published online: 4 February 2016

Abstract The distribution of the number of trials until the first k consecutive successes in a sequence of Bernoulli trials with success probability p is known as geometric distribution of order k. Let T_k be a random variable that follows a geometric distribution of order k, and $Y_1, Y_2, \ldots$ a sequence of independent and identically distributed discrete random variables which are independent of T_k. In the present article we develop some results on the distribution of the compound random variable $S_k = \sum_{t=1}^{T_k} Y_t$.

Keywords Compound distributions · Geometric distribution of order k · Phase-type distribution · Runs

Mathematics Subject Classification (2010) 60E05 · 60G40 · 60G50 · 60J10

1 Introduction

For a sequence $\xi_1, \xi_2, \cdots$ of binary trials each resulting in either a success “1” or a failure “0”, the number of trials until the first k consecutive successes can be formally defined as

$$T_k = \min\left(n : \xi_{n-k+1} = \ldots = \xi_n = 1\right).$$

✉ Markos V. Koutras
mkoutras@unipi.gr

Serkan Eryilmaz
serkan.eryilmaz@atilim.edu.tr

1 Department of Statistics and Insurance Science, University of Piraeus, Piraeus, Greece

2 Department of Industrial Engineering, Atilim University, İncek, Ankara, Turkey

If $\xi_1, \xi_2, \ldots$ are independent and identically distributed (iid) with $p = P(\xi_i = 1)$, then the distribution of the random variable T_k has been termed as geometric distribution of order k (see, e.g. Balakrishnan and Koutras (2002)). Distributional properties of the random variable T_k and several associated statistics (e.g. number of run occurrences, waiting times for the r-th occurrence of fixed length runs etc) have been extensively studied in the literature under the iid and Markov dependence assumption on the sequence of binary trials $\xi_1, \xi_2, \ldots$. The interested reader may consult, among others, Aki (1985), Aki and Hirano (1995), Koutras et al. (1995), Koutras (1997), Koutras and Papastavridis (1993), Philippou et al. (1983), Philippou and Makri (1986).

The geometric distribution of order k, besides its theoretical flavor, has been proved of substantial interest in numerous practical applications, ranging from quality control and reliability (a special reliability system named consecutive-k-out-of-n: F system is very closely related to this distribution), to non-parametric statistics, psychology, finance, ecology etc, Koutras and Alexandrou (1997), Eryilmaz (2010), Eryilmaz et al. (2011), Balakrishnan et al. (2009), Papastavridis and Koutras (1992),

Let us consider the random variable

$$S_k = \sum_{t=1}^{T_k} Y_t,$$

where $Y_1, Y_2, \ldots$ is a sequence positive valued iid random variables, independent of T_k. The compound random variable S_k is potentially useful for modeling the lifetime of a system under a particular run shock model. Consider a system that is subject to a sequence of shocks over time. Let Y_1 denote the time when the first shock occurs; the magnitude of the shock is also assumed as random and is described by a continuous random variable X_1. Moreover, denote by Y_t the interarrival time between the $(t-1)-th$ and $t-th$ shocks, and by X_t the respective magnitude of the $t-th$ shock, $t \geq 2$. Assume that our system fails if the magnitudes of the k consecutive shocks exceed a prespecified level $c > 0$. If the magnitudes of the shocks are independent, then the random variable T_k has a geometric distribution of order k with $p = P(\xi_t = 1) = P(X_t \geq c)$, and the compound random variable S_k represents the total lifetime of the system. As far as the shocks magnitude X_t are concerned they are usually modeled by phase-type distributions. For related reliability shock models, see, e.g. Sumita and Shanthikumar (1985), Gut (1990), Mallor and Omey (2001), Montoro-Cazorla et al. (2009).

The random variable S_k might also be useful in actuarial risk analysis. In this case the random variables X_t and Y_t respectively represent the claim size and the times between successive claims in a certain portfolio. Thus S_k defines the random time until the occurrence of k consecutive claims above a critical threshold.

It should be mentioned that, sums of independent random variables of the form

$$S = \sum_{t=1}^{N} Y_t$$

where the number of terms in N is itself a random variable, have attracted substantial research interest for many years. The most popular terminology for the distribution generated by such a random sum is *compound distribution*, while the terms *generalized distributions* and *stopped sum distributions* have also been in use at a smaller extend. Feller (1968), in his classic book on discrete probability, considered the special case where $Y_1, Y_2, \ldots$ is a sequence of integral-valued iid random variables, to illustrate the use of generating functions, while Charalambides (2005) presented a combinatorial approach to study the distribution of S. Both authors provided interesting applications of such random sums in several areas including ecology (population expansion), molecular biology, operations research (service times) etc. In these applications the random variable N is simply an enumerating variable (usually refers to the number of events in a specific time period); in our set-up N denotes waiting time for observing k consecutive occurrences of a specific event.

In the present article, we study the distribution of the random variable S_k when $Y_t's$ are discrete. We obtain recursive and nonrecursive formulae for the computation of the probability mass function (pmf) of the random variable S_k. More specifically, taking advantage of the fact that the distribution of the random variable T_k can be represented as a phase-type distribution, our derivations for finding the pmf of $S_k$are mainly based on the phase-type modeling of the random variable T_k.

A discrete phase-type distribution of order d is the distribution of the time T to absorption in a finite discrete time Markov chain with d transient states and one absorbing state, say "0". Let us denote by Λ_0 the $(d+1)\times(d+1)$ transition probability matrix of the Markov chain and by $\pi_0 = (\pi_1, \pi_2, \ldots, \pi_d, \pi_{d+1})'$ the respective initial probability (column) vector; from now on we assume that the absorption state of the Markov chain has been placed as the last state of the chain, i.e. state $d + 1$. Then, the probability mass function (pmf) of the discrete phase-type random variable T may be expressed in the form

$$P(T = t) = \pi'\Lambda^{t-1}\mathbf{u}, \quad t = 1, 2, \ldots \tag{1.1}$$

where Λ is the $d \times d$ substochastic matrix which includes the transition rates among the d transient states, $\pi = (\pi_1, \pi_2, \ldots, \pi_d)'$ is the initial transition probability vector with the entry corresponding to the absorption state removed and $\mathbf{u} = (I_d - \Lambda)\mathbf{1}$ is a column vector including all transition probabilities from the transient states to the absorbing state ($\mathbf{1} = (1, 1, \ldots, 1)'$ while I_d denotes the $d \times d$ identity matrix). For a discrete phase-type random variable T with pmf given by (1.1) we shall say that T follows a phase-type distribution of order d with parameters π, Λ, and we shall use the notation $T \sim PH_d(\pi, \Lambda)$.

For a detailed discussion of phase-type distributions and their properties, we refer to Neuts (1981) and He (2014).

The present paper is organized as follows. In Section 2, we provide some general results for the compound geometric distribution of order k. Explicit formulae are given for the probability generating function and moments of $S_k = \sum_{t=1}^{T_k} Y_t$ as well as a recursive scheme for the calculation of its probability mass function. A nonrecursive formula is also given for the probability mass function of $S_k = \sum_{t=1}^{T_k} Y_t$ when the common distribution of the random variables $Y_1, Y_2, \ldots$ is a phase-type distribution. In Section 3 we present several results in the special case when $Y_1, Y_2, \ldots$ have a geometric or negative binomial distribution. Section 4 deals with a generalization of the compound geometric distribution of order k, namely the compound negative binomial distribution of order k. In Section 5 we provide some numerical results for the computational effort required when applying each of the proposed approaches and a few comments on the shape of the probability mass function of the compound geometric distribution of order k. Finally, Section 6 contains some concluding remarks on the topic addressed in the present article.

2 General Results for the Compound Geometric Distribution of Order *k*

Let us start our study of the compound geometric distribution of order k by presenting some results related to the probability generating function (pgf) and moments of $S_k = \sum_{t=1}^{T_k} Y_t$.

We shall be denoting by

$$P_{S_k}(z)=E(z^{S_k})=\sum_{t=1}^{\infty} f_k(t)z^t=\sum_{t=1}^{\infty} P(S_k=t)z^t \text{and } P_{T_k}(z)=E(z^{T_k})=\sum_{t=1}^{\infty} P(T_k=t)z^t$$

the pgf's of T_k and S_k respectively, and by

$$P_Y(z)=E(z^{Y_t})=\sum_{x=1}^{\infty} P(Y_t=x)z^x$$

the common pgf's of all $Y_t's$, $t=1,2\ldots$. Applying the well known formula for the pgf of a random sum of random variables (see e.g. Bowers et al. (1997) or Feller (1968)) we may express the pgf of $S_k=\sum_{t=1}^{T_k} Y_t$ as follows

$$P_{S_k}(z)=P_k(P_Y(z)).$$

Making use of the following formula which gives the pgf of $P_{T_k}(z)$ (see e.g. Balakrishnan and Koutras (2002))

$$P_{T_k}(z)=E(z^{T_k})=\frac{(pz)^k-(pz)^{k+1}}{1-z+qp^kz^{k+1}}$$

we readily deduce the next expression for the pgf of the compound geometric distribution of order k

$$P_{S_k}(z)=E(z^{S_k})=\frac{(pP_Y(z))^k-(pP_Y(z))^{k+1}}{1-P_Y(z)+qp^k(P_Y(z))^{k+1}}. \tag{2.1}$$

The mean and variance of the compound geometric distribution of order k are given by the formulae

$$E(S_k) = \frac{1-p^k}{(1-p)p^k}E(Y_t)$$
$$Var(S_k) = \frac{1-p^k}{(1-p)p^k}Var(Y_t)+(E(Y_t))^2\frac{1-(2k+1)(1-p)p^k-p^{2k+1}}{(1-p)^2p^{2k}}.$$

This is an immediate consequence of the the well-known formulae for the mean and variance of random sum of variables, namely

$$E(S_k) = E\left(\sum_{t=1}^{T_k} Y_t\right)=E(T_k)E(Y_t),$$
$$Var(S_k) = Var\left(\sum_{t=1}^{T_k} Y_t\right)=E(T_k)Var(Y_t)+(E(Y_t))^2\,Var\,(T_k)\,.$$

and the following expressions for the mean and variance of the geometric distribution of order k (see, e.g. Balakrishnan and Koutras (2002))

$$E(T_k)=\frac{1-p^k}{(1-p)p^k},\quad Var(T_k)=\frac{1-(2k+1)(1-p)p^k-p^{2k+1}}{(1-p)^2p^{2k}}.$$

Formula (2.1) can be used, at least for some simple special cases, to get neat recurrence schemes for the pmf $f_k(t)=P(S_k=t), t=1,2,\ldots$ of S_k. As an illustration we mention that, if k =2 and $Y_1, Y_2,\ldots$ have geometric distribution with pmf

$$P\,(Y_i=y)=\theta\,(1-\theta)^{y-1}\,, y=1,2,\ldots$$

then $P_Y(z) = E(z^{Y_t}) = \theta z/(1-(1-\theta)z)$ and the pgf of the compound geometric distribution of order $k = 2$ reduces to

$$P_{S_2}(z) = E(z^{S_2}) = \frac{(p\theta z)^2}{1-[2-\theta(1+p)]z-[(1+p)\theta-p^2\theta^2-1]z^2}. \quad (2.2)$$

Therefore, the pmf $f_2(t) = P(S_2 = t)$, $t = 1, 2, \ldots$ obeys the next recursive scheme

$$f_2(t) = [2-\theta(1+p)]f_2(t-1) + [(1+p)\theta - p^2\theta^2 - 1]f_2(t-2), \quad t \geq 3$$

with initial conditions

$$f_2(2) = (p\theta)^2, \quad f_2(1) = 0.$$

It is noteworthy that, in cases where the pgf of the compound geometric distribution admits a simple closed expression, as the one shown above the method of partial fraction expansion can be used to derive an explicit formula for the pmf as well as an efficient asymptotic expression for it. The definite reference for this method is the classical Feller's (1968) book, while a nice application of this approach to run-related models applied to sampling inspection can be found in Shmueli and Cohen (2000).

For illustration purposes let us briefly present the application of the partial fraction expansion method for the special case of the compound geometric distribution of order $k = 2$ with $Y_1, Y_2, \ldots$ following the geometric distribution with parameter θ. To start with, let us write (2.2) in the form

$$P_{S_2}(z) = E(z^{S_2}) = c + \frac{u_1 z + u_0}{v_2 z^2 + v_1 z + v_0} = c + \frac{U(z)}{V(z)} \quad (2.3)$$

where

$$\begin{aligned} v_2 &= -2+\theta(1+p),\ v_1 = -(1+p)\theta + p^2\theta^2 + 1,\ v_0 = 1, \\ c &= p^2\theta^2/v_2, \\ u_1 &= -p^2\theta^2 v_1/v_2, \quad u_0 = -p^2\theta^2 v_0/v_2 \end{aligned}$$

and

$$U(z) = u_1 z + u_0, \quad V(z) = v_2 z^2 + v_1 z + v_0.$$

It is not difficult to verify that, for $p \neq 1$, the equation $V(z) = 0$ has 2 distinct real roots z_1, z_2 given by the formulae

$$z_{1,2} = \frac{-2+\theta(1+p) \pm \theta\sqrt{(1-p)(1+3p)}}{2[(1+p)\theta - p^2\theta^2 - 1]}.$$

Using the partial fraction expansion method for the ratio $U(z)/V(z)$ and adequate infinite geometric series, we can arrive at a power series expression of $P_{S_2}(z)$ from which the next simple formula arises for the pmf $f_2(t)$

$$f_2(t) = P(S_2 = t) = \frac{\rho_1}{z_1^{t+1}} + \frac{\rho_2}{z_2^{t+1}}, \quad t > 0. \quad (2.4)$$

The coefficients ρ_i, $i = 1, 2$ are given by the expression

$$\rho_i = -\frac{U(z_i)}{V'(z_i)} = -\frac{u_1 z_i + u_0}{2v_2 z_i + v_1}, \quad i = 1, 2.$$

Apparently, the evaluation of $f_2(t)$ by the aid of formula (2.4) is fairly easy for any t.

As stated in Feller (1968), in the general case the quantity $V(z)$ is a high degree polynomial and, although we can potentially develop an exact expression for the pmf, the labor involved in calculating the roots of the equation $V(z) = 0$ is usually prohibitive; this makes the method primarily of theoretical interest. Fortunately a simple and surprisingly good approximation can be derived for large t values by using only the root, say z_1 which is smaller in absolute value than all the other roots; more specifically, when $t \to \infty$, the following asymptotic formula holds true

$$f_2(t) = P(S_2 = t) \sim \frac{\rho_1}{z_1^{t+1}} \tag{2.5}$$

where the sign $\sim$ indicates that the ratio of the two sides tends to 1 as $t \to \infty$. In most cases this formula provides surprisingly good approximations even for relatively small values of t.

We shall now proceed to the development of some new formulae that exploit the theory of phase-type family of distributions. Eisele (2006) obtained recursive schemes for the pmf of the random variable $S = \sum_{t=1}^{T} Y_t$ when $Y_1, Y_2, \ldots$ is a sequence of positive valued iid random variables (discrete or continuous) with common pmf $f_Y(t)$, and T is a discrete random variable having a phase-type distribution of order d, say $T \sim PH_d(\pi, \Lambda)$. The recurrence schemes of Eisele (2006) are making use of two sets of coefficients that are computed from the substochastic $d \times d$ matrix Λ. The first set $b_1, \ldots, b_d$ is simply the set of coefficients of the characteristic polynomial of Λ, namely

$$\det(xI_d - \Lambda) = x^d + \sum_{i=1}^{d} b_i x^{d-i}.$$

The second sequence $a_1, \ldots, a_d$ is computed through $b_1, \ldots, b_d$ and $P(T = t)$, $t = 1, 2, \ldots, d$ by the following formulae

$$a_1 = P(T = 1), \qquad a_t = P(T = t) + \sum_{i=1}^{t-1} b_i P(T = t - i) \text{ for } i = 2, \ldots, d.$$

In order to apply Eisele's (2006) result for the compound geometric distribution of order k, we shall give first a Lemma providing the set of coefficients involved in his recursive scheme for the case of interest.

Lemma 1 *For the geometric distribution of order k we have,*

$$b_i = p^i - p^{i-1}, i = 1, \ldots, k \text{ and } a_i = \begin{cases} p^k, & \text{if } i = k \\ 0, & \text{otherwise.} \end{cases}$$

Proof As Tank and Eryilmaz (2014) have proved, the Geometric distribution of order k belongs to the family of phase-type distributions. More specifically, $T_k \sim PH_k(\pi, \Lambda)$ with $\pi = (1, 0, \ldots, 0)' = \mathbf{e}_1$ and

$$\Lambda = \begin{bmatrix} 1-p & p & 0 & \ldots & 0 \\ 1-p & 0 & p & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ 1-p & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{2.6}$$

□

It is not difficult to verify that, the characteristic polynomial of Λ is given by

$$\det(xI_k - \Lambda) = x^k + (p-1)x^{k-1} + \ldots + (p^{k-1} - p^{k-2})x + (p^k - p^{k-1}) = x^k + \sum_{i=1}^{k} b_i x^{k-i}.$$

Thus $b_i = p^i - p^{i-1}, i = 1, \ldots, k$ and taking into account that

$$P(T_k = i) = \begin{cases} 0 & \text{for } i = 1, 2, \ldots, k-1 \\ p^k & \text{for } i = k, \end{cases}$$

we can easily verify that

$$a_k = p^k \text{ and } a_i = 0 \text{ for } i = 1, 2, .., k-1.$$

We are now ready to establish an efficient set of recurrence relations for the compound geometric distribution of order k.

Proposition 1 *Assume that the support of the random variables* $Y_1, Y_2, \ldots$ *is* $\{y_0, y_0 + 1, ...\}$ *and denote by* $f_Y^{*j}(t)$ *the* $j - th$ *convolution of* $Y_i, Y_2, \ldots, Y_j$, *i.e.*

$$f_Y^{*j}(t) = P\left(\sum_{i=1}^{j} Y_i = t\right), \quad j = 1, 2, ... \quad . \tag{2.7}$$

Then the pmf $f_k(t) = P(S_k = t)$, $t = 1, 2, ...$ *of the respective compound geometric distribution of order k obeys the following recurrence scheme*

$$f_k(t) = p^k f_Y^{*k}(t) - \sum_{j=1}^{k} (p^j - p^{j-1}) \left(\sum_{u=1}^{t-1} f_k(u) f_Y^{*j}(t-u)\right) \text{for } t > y_0 k$$

with initial conditions

$$f_k(t) = 0 \text{ for } t < y_0 k, \qquad f_k(y_0 k) = [pP(Y = y_0)]^k.$$

Proof Eisele (2006) obtained the next recursion for the pmf of the random variable $S = \sum_{t=1}^{T} Y_t$ when T has a discrete phase-type distribution of order d and $Y_1, Y_2, \ldots$ is a sequence of positive valued iid random variables, independent of T:

$$P(S = t) = \sum_{j=1}^{\min(d,t)} a_j f_Y^{*j}(t) - \sum_{j=1}^{\min(d,t-1)} b_j \left(\sum_{u=1}^{t-1} P(S = u) f_Y^{*j}(t-u)\right), \text{ for } t \geq 1.$$

The result follows immediately by replacing the coefficients a_j, b_j by the expressions obtained in Lemma 1. □

It is of interest to note that, when the random variables $Y_1, Y_2, \ldots$ follow a phase-type distribution, one can evaluate the pmf $f_k(t) = P(S_k = t)$, $t = 1, 2, \ldots$ of the respective compound geometric distribution of order k by the aid of an exact formula similar to expression (1.1). This can be achieved by the following result.

Proposition 2 *Assume that the random variables* $Y_1, Y_2, \ldots$ *follow a phase-type distribution* $PH_c(\boldsymbol{\rho}, M)$ *of order c. Then the pmf* $f_k(t) = P(S_k = t)$, $t = 1, 2, \ldots$ *of the respective compound geometric distribution of order k is given by*

$$f_k(t) = P(S_k = t) = \boldsymbol{\sigma}' \Sigma^{t-1} (I_{ck} - \Sigma)\mathbf{1}, \quad t = 1, 2, \ldots \tag{2.8}$$

where

$$\boldsymbol{\sigma} = \boldsymbol{\rho} \otimes \boldsymbol{e}_1 (I_k - a\Lambda)^{-1}$$
$$\Sigma = M \otimes I_k + \boldsymbol{u}\boldsymbol{\rho}' \otimes (I_k - a\Lambda)^{-1} \Lambda,$$

Λ is the $k \times k$ transition probability matrix (2.6), $\boldsymbol{u} = (I_c - M)\boldsymbol{1}$ and $a = 1 - \boldsymbol{\rho}'1$. The notation $A \otimes B$represents the Kronecker product of two matrices A and B i.e. if $A = (a_{ij})_{n_1 \times n_2}$ then

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n_2}B \\ a_{21}B & a_{22}B & \dots & a_{2n_2}B \\ \vdots & & \vdots & \vdots \\ a_{n_1 1}B & a_{n_1 2}B & \dots & a_{n_1 n_2}B \end{bmatrix}.$$

Proof Conditioning on the random variable T_k we get

$$\begin{aligned} P(S_k = t) &= P\left(\sum_{i=1}^{T_k} Y_i = t\right) = \sum_{j=1}^{\infty} P\left(\sum_{i=1}^{T_k} Y_i = t | T_k = j\right) P(T_k = j) \\ &= \sum_{j=1}^{\infty} P\left(\sum_{i=1}^{j} Y_i = t | T_k = j\right) P(T_k = j) \end{aligned}$$

and making use of the independence between T_k and Y_i's we may express the pmf of S_k as follows

$$P(S_k = t) = \sum_{j=1}^{\infty} P\left(\sum_{i=1}^{j} Y_i = t\right) P(T_k = j) = \sum_{j=1}^{\infty} f_Y^{*j}(t) f_{T_k}(j).$$

The result can now be easily obtained by resorting to the next proposition given by Neuts (1981): If $\{f_N(\nu)\}$and $\{f_X(t)\}$ are the pmf's of two discrete PH-distributions$PH_d(\boldsymbol{\pi}, \Lambda)$, $PH_c(\boldsymbol{\rho}, M)$ of orders d and c respectively then the mixture

$$\sum_{\upsilon=0}^{\infty} f_N(\nu) f_X^{*\nu}(t)$$

follows a phase type distribution $PH_{cd}(\boldsymbol{\sigma}, \Sigma)$ of order cd, with parameters

$$\boldsymbol{\sigma} = \boldsymbol{\rho} \otimes \boldsymbol{\pi} (I_d - \alpha\Lambda)^{-1}, \quad \Sigma = M \otimes I_d + \boldsymbol{u}\boldsymbol{\rho}' \otimes (I_d - \alpha\Lambda)^{-1} \Lambda,$$

where $a = 1 - \boldsymbol{\rho}'\mathbf{1}$ and $\boldsymbol{u} = (I_c - M)\mathbf{1}$.

Since $T_k \sim PH_d(\boldsymbol{\pi}, \Lambda)$ with $d = k, \boldsymbol{\pi} = \mathbf{e}_1$ and Λ as given in expression (2.6), the proof is complete.

As an illustration of the application of Theorem 1, let us consider the case where the random variables $Y_1, Y_2, \dots$ follow a negative binomial distribution with parameters $r = 2$ and θ, i.e.

$$f_Y(y) = P(Y_i = y) = (y-1)\theta^2 (1-\theta)^{y-2}, \quad y = 2, 3, \dots .$$

The negative binomial distribution can be viewed as a phase type distribution $PH_c(\boldsymbol{\rho}, M)$ of order $c = 2$ with $\boldsymbol{\rho} = (1,\ 0)'$ and

$$M = \begin{bmatrix} 1-\theta & \theta \\ 0 & 1-\theta \end{bmatrix}.$$

Hence $a = 1 - \boldsymbol{\rho}'\mathbf{1} = 0$ and the rest quantities needed to apply Theorem 1 are easily computed as

$$\boldsymbol{\sigma} = \boldsymbol{\rho} \otimes \mathbf{e}_1 = (1, 0, 0, ..., 0)'_{1\times(2k)}, \quad \mathbf{u} = (I_2 - M)\mathbf{1} = (0, \theta)'$$

$$\Sigma = M \otimes I_k + \boldsymbol{u}\boldsymbol{\rho}' \otimes \Lambda = \begin{bmatrix} 1-\theta & \theta \\ 0 & 1-\theta \end{bmatrix} \otimes I_k + \begin{bmatrix} 1 & 0 \\ \theta & 0 \end{bmatrix} \otimes \Lambda = \begin{bmatrix} (1-\theta)I_k & \theta I_k \\ R & (1-\theta)I_k \end{bmatrix}$$

where

$$R = \begin{bmatrix} \theta(1-p) & \theta p & 0 & \dots & 0 \\ \theta(1-p) & 0 & \theta p & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \theta(1-p) & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Note now that $(I_{2k} - \Sigma)\mathbf{1} = (0, 0, ..., \theta p)'_{2k\times 1} = \theta p \boldsymbol{\tau}'$ where $\boldsymbol{\tau}' = (0, 0, ..., 1)$ is a unit vector of R^{2k} and replace all the above quantities in formula (2.7) to gain the following matrix representation for the pmf of the compound geometric distribution of order k

$$f_k(t) = P(S_k = t) = \boldsymbol{\sigma}'\Sigma^{t-1}(I_{2k} - \Sigma)\mathbf{1} = (\theta p)\boldsymbol{\sigma}' \begin{bmatrix} (1-\theta)I_k & \theta I_k \\ R & (1-\theta)I_k \end{bmatrix}^{t-1} \boldsymbol{\tau}', \ t = 1, 2, ...$$

On closing the present section we present a result pertaining to the stochastic behavior of the random variable $S_k$with respect to k. □

Proposition 3 *Let $Y_1, Y_2, \dots$ and $Y_1^*, Y_2^*, \dots$ be two sequences of independent random variable such that the terms of the first sequence are stochastically smaller than the respective terms of the second, i.e. $Y_t \leq_{st} Y_t^*$, $t = 1, 2, \dots$. If $k_1 \leq k_2$ then for the compound random variables*

$$S_{k_1} = \sum_{t=1}^{T_{k_1}} Y_t \text{ and } S_{k_2} = \sum_{t=1}^{T_{k_2}} Y_t^*$$

we have $S_{k_1} \leq_{st} S_{k_2}$.

Proof If $L_n$denotes the length of the longest run of successes (1's) in a sequence of iid binary trials $\xi_1, \xi_2, \dots, \xi_n$, then we have$P(T_k \geq n) = P(L_n \leq k)$. Thus for $k_1 \leq k_2$, we may state

$$P\left(T_{k_1} \geq n\right) = P\left(L_n \leq k_1\right) \leq P\left(L_n \leq k_2\right) = P\left(T_{k_2} \geq n\right)$$

which implies that $T_{k_1} \leq_{st} T_{k_2}$. The proof now follows by applying Theorem 1.A.4 of Shaked and Shanthikumar (1994). □

3 The Compound Geometric Distribution of Order *k* for Geometric and Negative Binomial *Y* 's

In the present section we shall apply the results established in Section 2 in the special case when $Y_1, Y_2, \dots$ have a negative binomial or geometric distribution. In Section 1 we presented an example from the reliability literature that motivated our research, where the random variables $Y_1, Y_2, \dots$ described the interarrival times between successive shocks. It si quite plausible to assume that the geometric and negative binomial distributions are suitable models for the interarrival times when we use discrete unit times (e.g. hours, days, weeks etc.). For example, if we assume that at each time unit, say day, a shock occurs with probability θ, then the number of days Y_1 until the first shock occurs will follow a geometric distribution with parameter θ; the same holds true for the interarrival times $Y_2, Y_3, \dots$ between the subsequent successive shocks. By a similar argument, we can substantiate the use of the negative binomial distribution for modeling the interarrival times if we assume that only after the occurrence of $r \geq 1$ shocks there is a positive probability that the shock magnitudes exceeds a prespecified level $c > 0$.

Let us now assume that $Y_1, Y_2, \ldots$ follow a negative binomial distribution with pmf

$$f_Y(y) = P\left(Y_i = y\right) = \binom{y-1}{r-1} \theta^r \left(1-\theta\right)^{y-r}, \quad y = r, r+1, \ldots \ .$$

Since the $j-th$ convolution of $Y_i, Y_2, ..., Y_j$ follows a negative binomial distribution with parameters rj and θ, i.e.

$$f_Y^{*j}(y) = P\left(\sum_{i=1}^{j} Y_i = y\right) = \binom{y-1}{rj-1} \theta^{rj} (1-\theta)^{y-rj}, \quad y = rj, rj+1, \ldots, 2, \ldots$$

a direct application of Proposition 1 yields the following recursive scheme for the pmf $f_k(t) = P(S_k = t), t = 1, 2, ...$ of the compound geometric distribution of order k.

$$\begin{aligned} f_k(t) &= \binom{t-1}{rk-1} (\theta^r p)^k (1-\theta)^{t-rk} \\ &\quad - \sum_{j=1}^{k} (p^j - p^{j-1}) \left(\sum_{u=1}^{t-rj} f_k(u) \binom{t-u-1}{rj-1} \theta^{rj} (1-\theta)^{t-u-rj} \right) \\ &\qquad \text{for } t > rk. \end{aligned}$$

The necessary initial conditions in order to launch the above scheme are

$$f_k(rk) = (\theta^r p)^k \text{ and } f_k(t) = 0 \text{ for } t < rk.$$

The pgf of the compound geometric distribution of order k with negative binomial compounding distribution can be easily deduced by replacing

$$P_Y(z) = \left(\frac{\theta z}{1-(1-\theta)z}\right)^r$$

in formula (2.1). More specifically we get

$$P_{S_k}(z) = E(z^{S_k}) = \frac{p^k (\theta z)^{rk} [(1-(1-\theta)z)^r - p(\theta z)^r]}{(1-(1-\theta)z)^{r(k+1)} - (\theta z)^r (1-(1-\theta)z)^{rk} + qp^k (\theta z)^{r(k+1)}}.$$

Let us next assume that $Y_1, Y_2, \ldots$ follow a geometric distribution with pmf

$$f_Y(y) = P\left(Y_i = y\right) = \theta \left(1-\theta\right)^{y-1}, \qquad y = 1, 2, \ldots \ .$$

Then one can get a recursive scheme and the pgf of the respective compound geometric distribution of order k by replacing $r = 1$ in the above formulae.

In this case, we can also develop a simple matrix expression for the direct evaluation of the pmf $f_k(t) = P(S_k = t), t = 1, 2, ...$ of the compound geometric distribution of order k. More spesifically, we have the following result.

Proposition 4 *If* $P\left(Y_i = y\right) = \theta\left(1-\theta\right)^{y-1}, y = 1, 2, \ldots$ *then the pmf of the compound geometric distribution of order k is given by the formula*

$$f_k(t) = P(S_k = t) = (\theta p) \mathbf{e}_1' \Sigma^{t-1} \mathbf{e}_k, \quad t = k, k+1, k+2, ...$$

where $\mathbf{e}_i, i = 1, 2, ..., k$ *denote the unit vectors of* R^k *and*

$$\Sigma = \begin{bmatrix} 1-\theta+\theta(1-p) & \theta p & 0 & 0 & \dots 0 & 0 \\ \theta(1-p) & 1-\theta & \theta p & 0 & \dots 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \theta(1-p) & 0 & 0 & 0 & 1-\theta & \theta p \\ \theta(1-p) & 0 & 0 & 0 & \dots 0 & 1-\theta \end{bmatrix}.$$

Proof The distribution of $Y_1, Y_2, \dots$ can be considered as a phase-type distribution $PH_c(\boldsymbol{\rho}, M)$ of order $c = 1$ with $\boldsymbol{\rho} = (1)_{1\times 1}$, $M = (1-\theta)_{1\times 1}$. Using these quantities, in the formulae appearing in Proposition 2 we get $a = 1 - \boldsymbol{\rho}'\mathbf{1} = 0$, $\boldsymbol{u} = (I_1 - M)\mathbf{1} = (\theta)_{1\times 1}$ and

$$\boldsymbol{\sigma} = \boldsymbol{\rho} \otimes \mathbf{e}_1 (I_k - a\Lambda)^{-1} = \boldsymbol{\rho} \otimes \mathbf{e}_1 = (1)_{1\times 1},$$

$$\Sigma = M \otimes I_k + \boldsymbol{u}\boldsymbol{\rho}' \otimes (I_k - a\Lambda)^{-1} \Lambda = (1-\theta) I_k + \theta\Lambda.$$

Replacing the $k \times k$ transition probability matrix Λ by formula (2.6) it is not difficult to verify that

$$\Sigma = (1-\theta)I_k + \theta\Lambda = \begin{bmatrix} 1-\theta+\theta(1-p) & \theta p & 0 & 0 & \dots 0 & 0 \\ \theta(1-p) & 1-\theta & \theta p & 0 & \dots 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \theta(1-p) & 0 & 0 & 0 & 1-\theta & \theta p \\ \theta(1-p) & 0 & 0 & 0 & \dots 0 & 1-\theta \end{bmatrix}$$

and $(I_k - \Sigma)\mathbf{1} = (\theta I_k - \theta\Lambda)\mathbf{1} = (\theta p)\mathbf{e}_k$. The result follows immediately by applying formula (2.8). □

Since the compound geometric distribution of order k with geometric compounding distributions can be regarded as a phase-type distribution $PH_k(\mathbf{e}_1', \Sigma)$ we can also use the following formula to compute its $m-th$ factorial moments (see Neuts (1981))

$$n_m = E[(S_k)_m] = E[S_k(S_k-1)\cdots(S_k-m+1)] = m!\mathbf{e}_1\Sigma^{m-1}(I-\Sigma)^{-m}\mathbf{1}, \quad m = 1, 2, ... \quad (3.1)$$

For example, in the special case $k = 2$, the $m-th$ factorial moments can be computed by

$$n_m = E[(S_2)_m] = m!(1,\ 0)\begin{bmatrix} 1-\theta p & \theta p \\ \theta(1-p) & 1-\theta \end{bmatrix}^{m-1} \begin{bmatrix} \theta p & -\theta p \\ -\theta(1-p) & \theta \end{bmatrix}^{-m} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

One could also manipulate in expression (3.1) to get a formula for the generating function of the factorial moments

$$M(z) = \sum_{m=1}^{\infty} n_m \frac{z^m}{m!},$$

as follows

$$M(z) = \sum_{m=1}^{\infty} (m!\mathbf{e}_1\Sigma^{m-1}(I-\Sigma)^{-m}\mathbf{1})\frac{z^m}{m!} = \mathbf{e}_1\left(\sum_{m=1}^{\infty}\Sigma^{m-1}(I-\Sigma)^{-m})z^m\right)\mathbf{1} =$$

$$= z\mathbf{e}_1\left(\sum_{m=1}^{\infty}(z\Sigma(I-\Sigma)^{-1})^{m-1}\right)(I-\Sigma)^{-1}\mathbf{1}$$

$$= z\mathbf{e}_1\left(I - z\Sigma(I-\Sigma)^{-1}\right)^{-1}(I-\Sigma)^{-1}\mathbf{1}.$$

Therefore

$$M(z)=\sum_{m=1}^{\infty}n_m\frac{z^m}{m!}=z\mathbf{e}_1\left((I-\Sigma)-z(I-\Sigma)\Sigma(I-\Sigma)^{-1}\right)^{-1}\mathbf{1}.$$

As an illustration, applying the last formula in the special case $k=2$, we readily get

$$M(z)=\sum_{m=1}^{\infty}n_m\frac{z^m}{m!}=\frac{(1+p)\theta z+[(1+p)\theta-1]z^2}{p^2\theta^2+\theta(2p^2\theta-p-1)z+[p^2\theta^2-(1+p)\theta+1]z^2}$$

from which the next fast recursive scheme arises for the computation of the factorial moments of the geometric distribution of order k with geometric compounding distributions

$$p^2\theta^2 n_m+\theta(2p^2\theta-p-1)n_{m-1}+[p^2\theta^2-(1+p)\theta+1]n_{m-2}=0 \qquad \text{for } m\geq 3.$$

4 A Generalization: Compound Negative Binomial Distribution of Order k

For a sequence $\xi_1,\xi_2,\ldots$ of binary trials, let $T_{r,k}$ denote the number of trials until the $r-th$ appearance of a success run of length k. If $\xi_1,\xi_2,\ldots$ are independent and identically distributed with $p=P(\xi_i=1)$, then the distribution of the random variable $T_{r,k}$ has been termed as type I negative binomial distribution of order k, see Balakrishnan and Koutras (2002).

In a similar fashion as before, if $Y_1,Y_2,\ldots$ is a sequence of positive valued iid random variables independent of $T_{r,k}$, we may name the distribution of the compound random variable

$$S_{r,k}=\sum_{t=1}^{T_{r,k}}Y_t$$

as compound negative binomial distribution of order k. This is apparently a generalization of the compound geometric distribution of order k ($S_{1,k}\stackrel{d}{=}S_k$).

Since $Y_1,Y_2,\ldots$ are iid random variables, $T_{r,k}$ can be represented as a sum of r independent random variables following a geometric distribution of order k; therefore $S_{r,k}$ can be considered as an r-fold convolution of the compound geometric distribution of order k, namely

$$S_{r,k}\stackrel{d}{=}S_k^{(1)}+\ldots+S_k^{(r)},$$

where $S_k^{(1)},\ldots,S_k^{(r)}$ are independent random variables each following a compound geometric distribution of order k.

If the random variables $Y_1,Y_2,\ldots$ belong to the phase-type family, we can readily conclude that the distribution of $S_{r,k}$ is a phase-type random variable as well. This conclusion arises immediately by the fact that, in this case, $S_k^{(1)},\ldots,S_k^{(r)}$ belong to the class of phase-type distributions and, as Neuts (1981) has proved, this family is closed under the convolution operator.

The distribution of the sum of two independent phase-type random variables has been established in Proposition 1.3.2 of He (2014). Using that result, we can proceed to the computation of the distribution of $S_{r,k}$ by an exact matrix-based formula as described in the following proposition.

Proposition 5 *Let $S_k^{(1)}, \ldots, S_k^{(r)}$ be independent random variables each having a compound geometric distribution of order k, with pmf of the form (2.8). Then the tail probability of the compound negative binomial distribution of order k can be expressed as*

$$P\left(S_{r,k} > t\right) = \pi_r' \Sigma_r^t \mathbf{1},$$

where $\boldsymbol{\pi}_r = (1, 0, \ldots, 0)'$ is a unit column vector with dimension rk, and $\Sigma_r$is a $rk \times rk$ matrix obtained recursively as follows

$$\Sigma_1 = \Sigma, \quad \Sigma_r = \begin{bmatrix} \Sigma & ((I_{ck} - \Sigma)\mathbf{1})\pi_{r-1}' \\ 0 & \Sigma_{r-1} \end{bmatrix}, for\ r = 2, 3, ...$$

As an illustration let us consider again the case where $Y_1, Y_2, \ldots$ have a geometric distribution with pmf

$$P(Y_i = y) = \theta(1-\theta)^{y-1}, \quad y = 1, 2, \ldots$$

and $k = 2$. Then, from Proposition 4, we have

$$\Sigma = \begin{bmatrix} 1-\theta p & \theta p \\ \theta(1-p) & 1-\theta \end{bmatrix},$$

and $(I_{ck} - \Sigma)\mathbf{1} = (I_2 - \Sigma)\mathbf{1} = (\theta p)\mathbf{e}_2$. For $i = 1, 2, 3$ we now deduce

$$\Sigma_1 = \Sigma = \begin{bmatrix} 1-\theta p & \theta p \\ \theta(1-p) & 1-\theta \end{bmatrix},$$

$$\Sigma_2 = \begin{bmatrix} \Sigma & (\theta p)\mathbf{e}_2\pi_1' \\ 0 & \Sigma_1 \end{bmatrix} = \begin{bmatrix} 1-\theta p & \theta p & 0 & 0 \\ \theta(1-p) & 1-\theta & \theta p & 0 \\ 0 & 0 & 1-\theta p & \theta p \\ 0 & 0 & \theta(1-p) & 1-\theta \end{bmatrix},$$

$$L_3 = \begin{bmatrix} \Sigma & (\theta p)\mathbf{e}_2\pi_2' \\ 0 & \Sigma_2 \end{bmatrix} = \begin{bmatrix} 1-\theta p & \theta p & 0 & 0 & 0 & 0 \\ \theta(1-p) & 1-\theta & \theta p & 0 & 0 & 0 \\ 0 & 0 & 1-\theta p & \theta p & 0 & 0 \\ 0 & 0 & \theta(1-p) & 1-\theta & \theta p & 0 \\ 0 & 0 & 0 & 0 & 1-\theta p & \theta p \\ 0 & 0 & 0 & 0 & \theta(1-p) & 1-\theta \end{bmatrix}$$

and the tail probability of the compound negative binomial distribution of order $k = 2$ can be evaluated by the formula

$$P\left(S_{3,2} > t\right) = \pi_3' \Sigma_3^t \mathbf{u}$$

where $\pi_3 = (1, 0, 0, 0, 0, 0)'$ and $\mathbf{u} = (1,\ 1,\ 1,\ 1,\ 1,\ 1)'$.

5 Numerical Results

Since in the present MS we have discussed more than one techniques for the evaluation of the probability mass function $f_k(t) = P(S_k = t)$, it is plausible to address the question

Table 1 CPU times (in seconds) by the recursive scheme and the matrix-based nonrecursive formula

k	t	Recursive method	Relative difference	Nonrecursive method	Relative difference
2	5	0.003154		0.000117	
	10	0.004658	0.476855	0.000118	0.008547
	20	0.010542	2.342422	0.000118	0.008547
	30	0.020530	5.509195	0.000118	0.008547
	50	0.052462	15.633481	0.000120	0.025641
	100	0.206116	64.350666	0.000120	0.025641
3	5	0.003169		0.000119	
	10	0.005380	0.697696	0.000119	0.000000
	20	0.014761	3.657936	0.000120	0.008403
	30	0.031114	8.818239	0.000122	0.025210
	50	0.084530	25.674030	0.000122	0.025210
	100	0.343953	107.536762	0.000122	0.025210
10	20	0.026238		0.000131	
	30	0.082149	2.130917	0.000136	0.038168
	50	0.272169	9.373085	0.000138	0.053435
	100	1.237378	47159767.275021	0.000141	0.076336
	300	16.36470	62370225.389207	0.000142	0.083969

which one is more efficient. In Table 1, we provide the CPU times needed for the computation of $f_k(t)$ by the aid of the recursive and the matrix-based nonrecursive schemes. The results refer to the special case of the compound geometric distribution of order $k$with $p = 0.5$ when the random variables $Y_1, Y_2, \ldots$ have geometric distribution with parameter $\theta = 0.8$. The computation was carried out for several values of k and t's ranging from small to sufficiently large values. The computer programs for both approaches were run in the MATLAB software on an Intel i7 (3.40 GHz) processor with 8GB of RAM. As easily conveyed from Table 1, there is a considerable difference in the CPU times between the two methods. For fixed k, the CPU time when working with the nonrecursive method is not substantially affected by the value of t; however, when working with the recursive scheme, for large values of t the CPU time increases significantly as compared to the small ones. This is more clear in columns 4 and 6 where have tabulated the relative change in CPU time with respect to time needed for the smallest t value ($t = 5$ for the first two blocks and $t = 20$ for the third one).

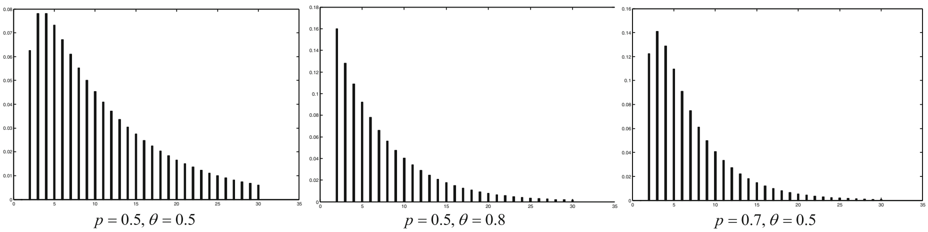
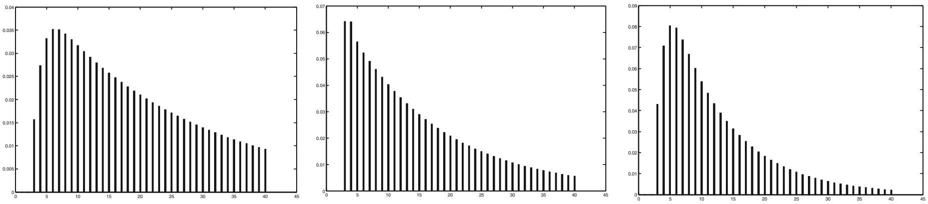


Fig. 1 Compound geometric distributions of order k =2



$p = 0.5, \theta = 0.5$ $\quad$ $p = 0.5, \theta = 0.8$ $\quad$ $p = 0.7, \theta = 0.5$

Fig. 2 Compound geometric distributions of order k =3

Should one be interested in the survival probability $P(S_k > t)$, the CPU times will become larger since, in this case, all $P(S_k = i), i = 1, \ldots, t$ should be computed in order to compute $P(S_k > t) = 1 - P(S_k \leq t)$. Although one might think that the recursive method would be more effective for this task (since all the values $P(S_k = i), i = 1, 2, \ldots, t-1$ have already been evaluated in order to get $P(S_k = t)$), this is not quite so because by the matrix (nonrecursive) approach the survival probability can be directly computed by the aid of Proposition 5.

The overall conclusion is that, the matrix-based formula seems to be more efficient if the random variables $Y_1, Y_2, \ldots$ have phase-type distribution, for both pmf and survival probability computations.

Needless to say, in simple cases where the partial fraction expansion method can be used for deriving the pmf from its probability generating function (see Section 2), the computations are by far faster than then ones compared above. Moreover, should one be interested in very large values of t, the corresponding asymptotic formula (2.5) is an excellent tool for a very fast and accurate computation of the pmf of the compound geometric distribution of order k.

In closing the present Paragraph we provide Figs. 1 and 2 where we have plotted the pmf of the compound geometric distribution of order k for k =2 and k =3 respectively in the special case when $Y_1, Y_2, \ldots$ have geometric distribution with pmf $P(Y_i = y) = \theta(1-\theta)^{y-1}, y = 1, 2, \ldots$.

Three different sets of values have been used for the parameters p and θ, namely $(p, \theta) = (0.5, 0.5), \ (0.5, 0.8), \ (0.7, 0.5)$. As made clear, the compound geometric distribution of order $k$is unimodal in all the graphs provided, and the tail of the distribution decays faster as p or θ increases while the reverse is observed when k increases. The analytical study of the behavior of the compound geometric distribution of order $k$as well as examination of statistical inference problems (e.g. parameter estimation and hypothesis testing based on independent samples from the distribution) will be the subject of a future work.

6 Conclusions

In this paper, we studied the distribution of the random sum

$$S_k = \sum_{t=1}^{T_k} Y_t$$

when T_k follows a geometric distribution of order k and $Y_1, Y_2, \ldots$ is a sequence of positive valued iid random variables, independent of T_k. This distribution has been termed as compound geometric distribution of order k.

We have obtained recursive and nonrecursive formulae for the evaluation of the probability mass function of S_k. The recursive equation given in Proposition 1 can be used for arbitrarily distributed discrete compounding variables while the nonrecursive matrix-based formula stated in Proposition 2, can only be used whenever the compounding variables follow a discrete phase-type distribution.

A numerical experimentation has also been carried out in order to illustrate the computational efficiency of the available alternative techniques for the evaluation of the probability mass function of the compound distribution.

Acknowledgments The authors would like to thank the two anonymous referees for their constructive comments that led to a significant improvement of the MS. Part of this work was carried out while S. Eryilmaz was visiting University of Piraeus. The research of M. V. Koutras has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Aristeia II. Investing in knowledge society through the European Social Fund.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

References

Aki S (1985) Discrete distributions of order k on a binary sequence. Ann Inst Stat Math 37:205–224

Aki S, Hirano K (1995) Joint distributions of numbers of success-runs and failures until the first consecutive k successes. Ann Inst Stat Math 47:225–235

Balakrishnan N, Koutras MV (2002) Runs and scans with applications. Wiley, New York

Balakrishnan N, Koutras MV, Bersimis S (2009) Run and frequency quota rules in process monitoring and acceptance sampling. J Qual Technol 41:66–81

Bowers NL, Hickman JC, Gerber HU, Nesbitt CJ, Jones DA (1997) Actuarial Mathematics, 2nd. Society of Actuaries

Charalambides ChA (2005) Combinatorial Methods in Discrete Distributions. Wiley, New York

Eisele K-T (2006) Recursions for compound phase distributions. Insurance: Math Econ 38:149–156

Eryilmaz S (2010) Review of recent advances in reliability of consecutive k-out-of-n and related systems. Proc. IMechE, Part O: J. Risk and Reliability 224:225–237

Eryilmaz S, Koutras MV, Triantafyllou IS (2011) Signature based analysis of m-consecutive-k-out-of-n: F Systems with exchangeable components. Nav Res Logist 58:344–354

Feller W (1968) An introduction to probability theory and its applications, vol I, 3rd edn. Wiley, New York

Gut A (1990) Cumulative shock models. Adv Appl Probab 22:504-507

He Q-M (2014) Fundamentals of matrix-analytic methods, Springer

Koutras MV (1997) Waiting time distributions associated with runs of fixed length in two-state Markov chains. Ann Inst Stat Math 49:123–139

Koutras MV, Alexandrou VA (1997) Non-parametric randomness tests based on success runs of fixed length. Stat Probability Lett 32:393–404

Koutras MV, Papadopoulos GK, Papastavridis SG (1995) Runs on a circle. J Appl Probab 32:396–404

Koutras MV, Papastavridis SG (1993) On the number of runs and related statistics. Statistica Sinica 3:277–294

Mallor F, Omey E (2001) Shocks, runs and random sums. J Appl Probab 38:438–448

Montoro-Cazorla D, Ocon RP, Segovia MC (2009) Shock and wear models under policy N using phase-type distributions. Appl Math Model 33:543–554

Neuts MF (1981) Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach. The Johns Hopkins University Press, Baltimore

Papastavridis SG, Koutras MV (1992) Consecutive k-out-of-n systems with maintainance. Ann Inst Stat Math 44:605–612

Philippou AN, Georghiou C, Philippou GN (1983) A generalized geometric distribution and some of its properties. Stat Probability Lett 1:171–175

Philippou AN, Makri FS (1986) Successes, runs and longest runs. Stat Probability Lett 4:211–215
Shaked M, Shanthikumar JG (1994) Stochastic orders and their applications, Academic Press
Shmueli G, Cohen A (2000) Run-related probability functions applied to sampling inspection. Technometrics 42:188–202
Sumita U, Shanthikumar JG (1985) A class of correlated cumulative shock models. Adv Appl Probab 17:344–366
Tank F, Eryilmaz S (2014) The distributions of sum, minima and maxima of generalized geometric random variables. Statistical Papers 56:1191–1203