Some Generalizations of Geometric Distribution in Bernoulli Trials by TPFG Methods

JIGUANG SHAO
Beijing Jiaotong University
Department of Mathematics
Shangyuancun 3, Haidian, Beijing, 100044
CHINA
jgshao@bjtu.edu.cn

SHENG FU
University of Chinese Academy of Sciences
School of Mathematical Sciences
Yuquanlu 19, Shijingshan, Beijing, 100049
CHINA
fusheng13@mails.ucas.ac.cn

Abstract: In this paper, by the transition probability flow graphs theory, we obtain the transition probability flow graphs of the random variables distributed as $G_k(p)$ and $NB_k(r,p)$, the generalizations of usual geometric distribution. We derive their probability generating functions, means, variances and their exact probability distributions, in addition, we reveal the correlation between them. For success runs in n Bernoulli trials, we find the new concise formulae for the probability distributions of L_n and $N_n^{(k)}$. Finally, proceed from the parameter $n = 2, 3, 4$, we derive the generalized geometric distribution denoted by $G(p_1, \cdots, p_n)$ and discuss its properties base on multivariate transition probability flow graphs methods.

Key–Words: geometric distribution; negative binomial distribution; success run; TPFG methods; probability generating function; probability distribution

1 Introduction

Philippou [16] defined the geometric distribution of order k and the negative binomial distribution of order k, Philippou and Makri [17] defined the binomial distribution of order k. Muselli [15] derived some probability distributions formulae in Bernoulli trials. The present paper is partly a continuation of their works. All the distributions in section 2 and 3 can be defined by success run, one of the important tools in the study of Bernoulli trials. There are several definitions about run (see, Han [11] and Schwager [18]), for example, we can define a success run to be a specified sequence of k consecutive success that may occur at some point in the series of Bernoulli trials, where k is the length of the run. In section 4, inspired by [3], [9] and [19], we define a new generalization geometric distribution in independent trials, and discuss its some probability properties. In the present paper, we try to show the transition probability flow graphs (TPFG) methods to study the distributions.

TPFG theory is a forceful tool for discussing some complicated discrete random variables. By decomposing the Markov chain formed by the variation of a nonnegative integer-valued random variable, ascertaining the states and routes, and setting probability functions to the routes, we get a flow graph of the process being similar to the transition probability graph of the chain. Based on the series-parallel operation rules, we can get the probability generating function of the random variable from the flow graph. The TPFG's prototype is the signal flow graphs theory applied to systems engineering widely, which was given by Mason [14]. Koyama [12] firstly introduced it into the study of sampling system. With the development of sampling inspection, Fan (1998) did lots of work for TPFG such that it gradually became a complete theory. We give a brief description for it as follows, for detail, the readers are referred to Fan [4, 5, 6].

Let τ be a nonnegative integer-valued random variable with probability space $(\Omega, \mathcal{F}, P)$, set $B_n = \{\tau = n\}$, for $B \in \mathcal{F}$, the transition probability function of τ is defined by $G_\tau(B;x) = \sum_{n=0}^{\infty} P(BB_n)x^n, |x| \leq 1$. When $B = \Omega$ we get the probability generating function of τ as $G_\tau(x) = \sum_{n=0}^{\infty} P(\tau = n)x^n$.

Consider a Markov chain that takes on countable number of possible values. The transition process from state A into B denoted by $R : A \to B$ is called a route, and its transition time named step is a random variable. By the Markov property that the steps of $A \to B$ and $B \to C$ are independent. The transition probability function of the route R is defined by $G_R(x) = \sum_{n=0}^{\infty} P_R(n)x^n$, where $P_R(n)$ is the $n-$step transition probability of R. The route from A into C by way of B denoted by $R_1 \cdot R_2$ is called a series route if the routes $R_1 : A \to B$ and $R_2 : B \to C$ are independent. The route denoted by $R_1 + R_2$ is called a parallel route of $R_1 : A \to B$ and $R_2 : A \to B$ if they are mutually exclusive. Then we have

Lemma 1 *Let $G_{R_1}(x)$ and $G_{R_2}(x)$ be respectively the transition probability functions of R_1 and R_2, then $G_{R_1 \cdot R_2}(x) = G_{R_1}(x) \cdot G_{R_2}(x)$ and $G_{R_1+R_2}(x) = G_{R_1}(x) + G_{R_2}(x)$.*

The route $A \to B (A \neq B)$ denoted by L is called a straight route if no state is repeated in it. The route denoted by C is called a loop route on state A if A can be repeated infinitely, and all the repeated routes are independent identically distributed. For the straight route and loop route, we have

Lemma 2 *The series-parallel connection rules of straight routes and loop routes are shown in Figure 1.*

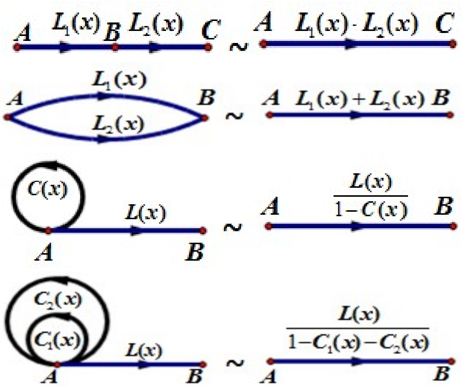

Figure 1: Connection rules for the straight and loop routes

Let $\tau_1, \tau_2, \cdots, \tau_n$ be n random variables, their joint probability space is $(\Omega, \mathcal{F}, P)$, for $A \in \mathcal{F}$, the joint transition probability function of $(\tau_1, \tau_2, \cdots, \tau_n)$ is given by

$$\begin{aligned} &G(x_1, x_2, \cdots, x_n; A) \\ &= \sum_{i_1, \cdots, i_n} P(\tau_1 = i_1, \cdots, \tau_n = i_n; A) x_1^{i_1} \cdots x_n^{i_n}, \end{aligned}$$

where $|x_k| \leq 1, k = 1, \cdots, n$. When $A = \Omega$, $G(x_1, \cdots, x_n; \Omega)$ is called the joint probability generating function of $(\tau_1, \cdots, \tau_n)$, denoted by $G(x_1, \cdots, x_n)$.

Lemma 3 *Let $G(x_1, x_2, \cdots, x_n)$ be the joint probability generating function of $(\tau_1, \tau_2, \cdots, \tau_n)$, then $G(1, \cdots, 1, x_k, 1, \cdots, 1)$ is the probability generating function of τ_k, $k = 1, 2, \cdots, n$, and $G(x, x, \cdots, x)$ is the probability generating function of $\tau = \tau_1 + \cdots + \tau_n$, which does not depend on the independence of $\tau_1, \tau_2, \cdots, \tau_n$.*

Lemma 4 *Let $G_\eta(x)$ be the probability generating function of η, then*

$$E\eta = G'_\eta(1),$$

$$Var\eta = G''_\eta(1) + G'_\eta(1) - G'^2_\eta(1).$$

2 The geometric distribution of order k

In this section, we shall generalize the usual geometric distribution into the geometric distribution of order k by TPFG methods, furthermore, we will discuss its properties.

Let $\xi_{(k)}$ be the number of trials until the occurrence of the success run with length k in Bernoulli trials with success probability p. We denote the probability distribution of $\xi_{(k)}$ by $G_k(p)$ and call it the geometric distribution of order k with parameter p. Then we have

Theorem 5 *The mean and variance of $\xi_{(k)}$, denoted by $E\xi_{(k)}$ and $Var\xi_{(k)}$, are given by*

$$E\xi_{(k)} = \frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^k},$$

$$Var\xi_{(k)} = \frac{q\left[\sum_{l=0}^{k-1} \binom{l+2}{2} p^l + \sum_{l=k}^{2k-2} \binom{2k-l}{2} p^l\right]}{p^{2k}}.$$

Proof: Let τ_n be a state that the trial process is in at the end of the foregoing n trials, where the number of trials n is also named transition time. then $\{\tau_n, n = 1, 2, \cdots\}$ is a Markov chain with state space $S = \{0, 1, 2, \cdots, k\}$, where 0 is the beginning state B, and k is the ending state E. The event $\{\tau_n = s, s \in S\}$ denotes the occurrence of s consecutive successes at the end of the foregoing n trials. Hence, $\{\tau_n = k\} = \{\xi_{(k)} = n\}$. From the beginning to the ending, we can obtain the transition probability flow graphs of $\{\tau_n, n = 1, 2, \cdots\}$, i.e., the TPFG of $G_k(p)$ as shown in Figure 2.

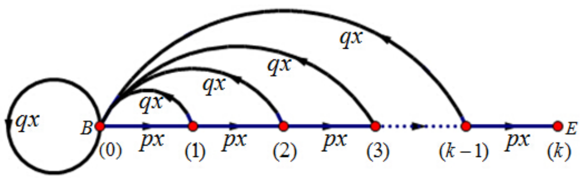

Figure 2: The TPFG of $G_k(p)$

There are k parallel loop routes at state B with respectively the transition probability functions $qx, qx(px), qx(px)^2, \cdots, qx(px)^{k-1}$, and k series straight routes from B to E with the same transition probability function px. Then we get the transition probability functions of the loop route and the straight route from B to E denoted respectively by $C(x)$ and $L(x)$ as follows

$$C(x) = \sum_{l=0}^{k-1}(qx)(px)^l = \frac{qx[1-(px)^k]}{1-px},$$
$$L(x) = (px)\cdot(px)\cdots(px) = (px)^k.$$

By Figure 2 and Lemma 2, we get the probability generating function of $\xi_{(k)}$ (the transition time of the route $B \to E$)

$$G_{\xi_{(k)}}(x) = \frac{L(x)}{1-C(x)} = \frac{p^k x^k(1-px)}{1-x+qp^kx^{k+1}}. \qquad (1)$$

To derive the mean and variance of $\xi_{(k)}$, we differentiate $G_{\xi_{(k)}}(x)$ and evaluate at $x=1$, that is

$$G'_{\xi_{(k)}}(1) = \frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^k},$$
$$G''_{\xi_{(k)}}(1) = \frac{q[1-p^k-(k+1)qp^k+qp^{2k}]}{q^2p^{2k}}.$$

The results follow from Lemma 4. This completes the proof. □

Remark 6 *1) When $p=1/2$, $E\xi_{(k)} = 2(2^k-1)$ has been obtained by Barry [2].*

2) The equation (1) has also been obtained by Feller [7], Philippou [16] and Aki [1].

After that, we expand the probability generating function $G_{\xi_{(k)}}(x)$ into power series for deriving the exact probability distribution of $\xi_{(k)}$.

$$G_{\xi_{(k)}}(x) = p^kx^k(1-px)\sum_{n=0}^{\infty}(x-qp^kx^{k+1})^n$$
$$= p^kx^k(1-px)\sum_{n=0}^{\infty}\sum_{m=0}^{n}\binom{n}{m}(-qp^k)^m x^{mk+n},$$

where

$$\sum_{n=0}^{\infty}\sum_{m=0}^{n}\binom{n}{m}(-qp^k)^m x^{mk+n} = \sum_{j=0}^{k}\binom{j}{0}x^j$$
$$+\sum_{j=1}^{k+1}\left[\binom{j}{1}(-qp^k)+\binom{k+j}{0}\right]x^{k+j}$$
$$+\sum_{j=2}^{k+2}\left[\binom{j}{2}(-qp^k)^2+\binom{k+j}{1}(-qp^k)\right.$$
$$\left.+\binom{2k+j}{0}\right]x^{2k+j}$$
$$+\sum_{j=3}^{k+3}\left[\binom{j}{3}(-qp^k)^3+\binom{k+j}{2}(-qp^k)^2\right.$$
$$\left.+\binom{2k+j}{1}(-qp^k)+\binom{3k+j}{0}\right]x^{3k+j}+\cdots$$
$$+\sum_{j=l}^{k+l}\left[\sum_{m=0}^{l}\binom{mk+j}{l-m}(-qp^k)^{l-m}\right]x^{lk+j}+\cdots$$
$$=\sum_{n=0}^{\infty}a_n^{(k)}x^n,$$

where $a_n^{(k)}$ is given by

$$a_n^{(k)} = \sum_{j=0}^{[\frac{n}{k+1}]}\binom{n-jk}{j}(-qp^k)^j. \qquad (2)$$

It follows that

$$G_{\xi_{(k)}}(x) = (p^kx^k-p^{k+1}x^{k+1})\sum_{n=0}^{\infty}a_n^{(k)}x^n$$
$$= p^kx^k+\sum_{n=k+1}^{\infty}\left(p^ka_{n-k}^{(k)}-p^{k+1}a_{n-k-1}^{(k)}\right)x^n.$$

Thus, we obtain the following result:

Theorem 7 *The probability distribution of the kth order geometric variable $\xi_{(k)}$ is*

$$P(\xi_{(k)}=n) = p^ka_{n-k}^{(k)}-p^{k+1}a_{n-k-1}^{(k)}, \qquad (3)$$

where $a_{n-k}^{(k)}$, $a_{n-k-1}^{(k)}$ can be derived from equation (2) and $a_n^{(k)}\equiv 0$ if $n<0$.

Corollary 8 *For $k=1$, $\xi_{(1)}$ is a random variable distributed as usual geometric distribution.*

Proof: By equations (2) and (3), we have

$$P\left(\xi_{(1)}=n\right) = p\left(a_{n-1}^{(1)}-pa_{n-2}^{(1)}\right)$$
$$= p\sum_{j=0}^{[\frac{n-1}{2}]}\binom{n-j-1}{j}(-qp)^j$$
$$-p^2\sum_{j=0}^{[\frac{n-2}{2}]}\binom{n-j-2}{j}(-qp)^j$$
$$= p\sum_{j=0}^{n-1}\binom{n-1}{j}(-p)^j$$
$$= p\,(1-p)^{n-1} = q^{n-1}p,$$

where we used equation $\binom{m-1}{l-1}+\binom{m-1}{l}=\binom{m}{l}$. Corollary 8 has been proven. □

Remark 9 *It follows from Corollary 8 that $G_k(p)$ is a generalized geometric distribution.*

Let L_n be the length of the longest success run in n Bernoulli trials with success probability p and set $R_n(k) = P(L_n < k)$. The event $\{\xi_{(k)} = n\}$ is equivalent to that the last k trials are all successful in n trials, the $(n-k)$th trial is fail and all the lengths of the success runs are less than k in the foregoing $(n-k-1)$ trials. Therefore

$$P(\xi_{(k)} = n) = qp^k R_{n-k-1}(k). \tag{4}$$

By considering equations (2), (3) and (4), we have

$$\begin{aligned} R_n(k) &= q^{-1}p^{-k} \cdot P(\xi_{(k)} = n+k+1) \\ &= \frac{1}{q}\sum_{j=0}^{[\frac{n+1}{k+1}]} \binom{n-jk+1}{j} (-qp^k)^j \\ &- \frac{p}{q}\sum_{j=0}^{[\frac{n}{k+1}]} \binom{n-jk}{j} (-qp^k)^j. \end{aligned} \tag{5}$$

Then we have

Corollary 10 *The distribution of the longest success run L_n in n Bernoulli trials is*

$$P(L_n = k) = R_n(k+1) - R_n(k),$$

where $R_n(k+1)$, $R_n(k)$ can be derived from (5).

Remark 11 *Corollary 10 is more concise than the same work of Fu [8] and Muselli [15].*

3 The negative binomial distribution of order k

In this section, we shall derive the probability generating function of the negative binomial distribution of order k, and then discuss its relationship with some other distributions.

Let $\xi_{(k,r)}$ be a random variable denoting the number of trials until the rth occurrence of the success run with length k in Bernoulli trials with success probability p. We denote the probability distribution of $\xi_{(k,r)}$ by $NB_k(r,p)$ and call it the negative binomial distribution of order k with parameter vector (r,p).

Theorem 12 *The mean and variance of $\xi_{(k,r)}$ are given by*

$$E\xi_{(k,r)} = \frac{r}{p} + \frac{r}{p^2} + \cdots + \frac{r}{p^k},$$

$$Var\xi_{(k,r)} = \frac{rq\left[\sum_{l=0}^{k-1}\binom{l+2}{2}p^l + \sum_{l=k}^{2k-2}\binom{2k-l}{2}p^l\right]}{p^{2k}}.$$

Proof: Similar to the proof of Theorem 5, let $\{\tau_n, n = 1, 2, \cdots\}$ be a Markov chain with state space $S = \{0, 1, 2, \cdots, r\}$, where n is the number of trials. The event $\{\tau_n = s, s \in S\}$ denotes the sth occurrence of the success run with length k at the end of the foregoing n trials. Hence, $\{\tau_n = r\} = \{\xi_{(k,r)} = n\}$. Thus we have the transition probability flow graphs of $\{\tau_n, n = 1, 2, \cdots\}$ as shown in Figure 3, where $k = 2$ for convenience.

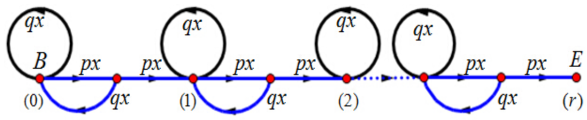

Figure 3: The TPFG of $NB_k(r,p)$

We may employ Figure 3 and Lemma 2 to get the probability generating function of $\xi_{(k,r)}$

$$G_{\xi_{(k,r)}}(x) = \left(\frac{p^k x^k - p^{k+1}x^{k+1}}{1 - x + qp^k x^{k+1}}\right)^r. \tag{6}$$

By Lemma 4, we can derive the mean and variance of $\xi_{(k,r)}$. The proof is complete. □

Remark 13 *Following the equations (1) and (6), if $X_i (i = 1, \cdots, r)$ are independently distributed as $G_k(p)$, then $X_1 + \cdots + X_r$ is distributed as $NB_k(r,p)$.*

Theorem 14 *The probability distribution of $\xi_{(k,r)}$ is*

$$P(\xi_{(k,r)} = n) = \sum_{l=0}^{r} \binom{r}{l} (-p)^l a_{n-kr-l}^{(k,r)} p^{kr}, \tag{7}$$

where

$$a_n^{(k,r)} = \sum_{j=0}^{[\frac{n}{k+1}]} \binom{r+n-jk-1}{r-1, j, n-jk-j} (-qp^k)^j \tag{8}$$

and $a_n^{(k,r)} \equiv 0$ if $n < 0$.

Proof:

$$G_{\xi_{(k,r)}}(x) = p^{kr}x^{kr}(1-px)^r \left[1 - (x - qp^k x^{k+1})\right]^{-r},$$

where

$$\begin{aligned} \left[1 - (x - qp^k x^{k+1})\right]^{-r} &= 1 + \sum_{j=1}^{k} \binom{r+j-1}{j} x^j \\ + \sum_{j=1}^{k+1} &\left[\binom{r+j-1}{j}\binom{j}{1}(-qp^k)\right. \\ + &\left.\binom{r+k+j-1}{k+j}\binom{k+j}{0}\right] x^{k+j} + \cdots \end{aligned}$$

$$+\sum_{j=l}^{k+l}\left[\sum_{m=0}^{l}\binom{r+mk+j-1}{mk+j}\binom{mk+j}{l-m}\right.$$
$$\left.\times(-qp^k)^{l-m}\right]x^{lk+j}+\cdots=\sum_{n=0}^{\infty}a_n^{(k,r)}x^n,$$

where $a_n^{(k,r)}$ is given by (8). Therefore

$$G_{\xi_{(k,r)}}(x)=p^{kr}x^{kr}(1-px)^r\sum_{n=0}^{\infty}a_n^{(k,r)}x^n$$
$$=\sum_{n=0}^{r-1}\left(\sum_{l=0}^{n}\binom{r}{l}(-p)^l a_{n-l}^{(k,r)}\right)p^{kr}x^{kr+n}$$
$$+\sum_{n=r}^{\infty}\left(\sum_{l=0}^{r}\binom{r}{l}(-p)^l a_{n-l}^{(k,r)}\right)p^{kr}x^{kr+n}$$
$$=\sum_{n=kr}^{kr+r-1}\left(\sum_{l=0}^{n-kr}\binom{r}{l}(-p)^l a_{n-kr-l}^{(k,r)}\right)p^{kr}x^n$$
$$+\sum_{n=kr+r}^{\infty}\left(\sum_{l=0}^{r}\binom{r}{l}(-p)^l a_{n-kr-l}^{(k,r)}\right)p^{kr}x^n.$$

From the above we can get the probability distribution of $\xi_{(k,r)}$. □

Corollary 15 *$NB_1(r,p)$ is a negative binomial distribution with parameter vector (r,p).*

Proof: From equations (7) and (8), we have

$$P(\xi_{(1,r)}=n)=\sum_{l=0}^{r}\binom{r}{l}(-p)^l a_{n-r-l}^{(1,r)}p^r$$
$$=\sum_{l=0}^{r}\binom{r}{l}(-p)^l\sum_{j=0}^{[\frac{n-r-l}{2}]}\binom{n-l-j-1}{r-1}$$
$$\times\binom{n-r-l-j}{j}(-qp)^j p^r$$
$$=\sum_{l=0}^{r}\binom{r}{l}\sum_{j=0}^{[\frac{n-r-l}{2}]}\sum_{s=0}^{j}\binom{n-l-j-1}{r-1}$$
$$\times\binom{n-r-l-j}{j}\binom{j}{s}(-p)^{l+s+j}p^r$$
$$=\sum_{l=0}^{r}\sum_{j=0}^{[\frac{n-r-l}{2}]}\sum_{s=0}^{j}\binom{r}{l}\binom{n-l-j-1}{r-1}$$
$$\times\binom{n-r-l-j}{j}\binom{j}{s}(-p)^{l+s+j}p^r$$
$$=\binom{n-1}{r-1}\sum_{j=0}^{n-r}(-p)^j p^r$$
$$=\binom{n-1}{r-1}(1-p)^{n-r}p^r. \qquad (9)$$

The result is proven. □

Remark 16 *The proof for Corollary 15 can verify the correctness of equation (7).*

Moreover, following from

$$P(\eta=n)=\binom{n-1}{r-1}(1-p)^{n-r}p^r,$$

the formula for the probability distribution of the negative binomial variable η, we have its probability generating function

$$G_\eta(x)=\sum_{n=r}^{\infty}\binom{n-1}{r-1}(1-p)^{n-r}p^r x^n$$
$$=(px)^r\sum_{n=r}^{\infty}\binom{n-1}{r-1}(qx)^{n-r}$$
$$=(px)^r\sum_{k=0}^{\infty}\binom{r-1+k}{r-1}(qx)^k=\left(\frac{px}{1-qx}\right)^r,$$

on the other hand,

$$G_{\xi_{(1,r)}}(x)=\left(\frac{px-p^2x^2}{1-x+qpx^2}\right)^r=\left(\frac{px}{1-qx}\right)^r,$$

which illustrates that $NB_1(r,p)$ is a negative binomial distribution.

Corollary 17 $NB_k(1,p)=G_k(p)$.

Corollary 18 *$NB_1(1,p)$ is the usual geometric distribution.*

Theorem 19 *Let $N_n^{(k)}$ be the number of success run with length k in n Bernoulli trials with success probability p. The probability distribution of $N_n^{(k)}$ denoted by $B_k(n,p)$ is called the binomial distribution of order k with parameter vector (n,p). We have*

$$P(N_n^{(k)}=r)=\sum_{s=0}^{k-1}\sum_{l=0}^{r+1}\tbinom{r+1}{l}(-p)^l a_{n-kr-l-s}^{(k,r+1)}p^{kr+s},$$

where $a_{n-kr-l-s}^{(k,r+1)}$ can be derived from (8).

Proof: For $\xi_{(k,r+1)}$ is distributed as $NB_k(r+1,p)$, let $C_k=\{\xi_{(k,r+1)}=n+k-s|(k-s)\ successes\}$ denote the event that $(r+1)$ success runs of length k occur in $(n+k-s)$ trials and the posterior $(k-s)$ successes are deleted, where $s=0,1,\cdots,k-1$. We shall find that $\cup_{s=0}^{k-1}C_k$ means all the possible ways the r success runs occur in n Bernoulli trials, i.e., $\{N_n^{(k)}=r\}=\cup_{s=0}^{k-1}C_k$. Hence we have

$$P(N_n^{(k)}=r)=P(\cup_{s=0}^{k-1}C_k)$$

$$= P(\bigcup_{s=0}^{k-1}\{\xi_{(k,r+1)} = n+k-s|(k-s)\ successes\})$$
$$= \sum_{s=0}^{k-1} p^{-(k-s)} \cdot P(\xi_{(k,r+1)} = n+k-s). \quad (10)$$

By Theorem 14, we can get the probability distribution of $N_n^{(k)}$. □

Remark 20 *A equivalent formula of Theorem 19 has also been obtained by Godbole [10].*

Corollary 21 *For $k = 1$, $N_n^{(1)}$ is distributed as the usual binomial distribution.*

Proof: By considering (9) and (10), we have

$$P\left(N_n^{(1)} = r\right) = p^{-1} \cdot P(\xi_{(1,r+1)} = n+1)$$
$$= p^{-1}\binom{n+1-1}{r+1-1}(1-p)^{(n+1)-(r+1)}p^{r+1}$$
$$= \binom{n}{r}(1-p)^{n-r}p^r.$$

The proof of Corollary 21 is complete. □

4 Some properties of $G(p_1, \cdots, p_n)$

In sections 2 and 3, base on Bernoulli trials, we generalize the usual geometric distribution into $G_k(p)$ and $NB_k(r,p)$, moreover, we show the correlation between them. In this section, we shall discuss a new generalized geometric distribution denoted by $G(p_1, \cdots, p_n)$ in independent trials. And the corresponding methods named multivariate transition probability flow graphs.

Suppose that each independent trial result is one of the events A_1, A_2,$\cdots$, A_n, which are mutually exclusive with respective success probability $p_1, p_2, \cdots, p_n$, such that $\sum_1^n p_k = 1$. Obviously, the independent trial is a generalization of Bernoulli trial. Let ξ be the number of trials required until every $A_k, k = 1, 2, \cdots, n$ has exactly occurred, then $\xi = \xi_1 + \xi_2 + \cdots + \xi_n$ is called the generalized geometric variable with parameter vector $(p_1, p_2, \cdots, p_n)$, denoted by $G(p_1, p_2, \cdots, p_n)$, where ξ_k is the occurrence number of A_k, $k = 1, 2, \cdots, n$.

Theorem 22 *Let ξ be the generalized geometric random variable with parameter vector (p_1, p_2), then*

$$E\xi = \frac{1}{p_1} + \frac{1}{p_2} - 1, \quad (11)$$

$$Var\xi = \frac{p_1^2}{p_2^2} + \frac{p_2^2}{p_1^2} + \frac{p_1}{p_2} + \frac{p_2}{p_1} - 2, \quad (12)$$

$$P(\xi = n) = p_1p_2(p_1^{n-2} + p_2^{n-2}), n = 2, 3, \cdots. \quad (13)$$

Proof: Starting from the state B, the process comes to the next state as soon as a new event occurs in a trial, otherwise, it keeps stopping in the original state. Let $E_{12} = E_{21} = E$ denote the ending state, the route $B \to E$ constitutes a Markov chain as shown in Figure 4, where $l_1 = c_1 = p_1x_1$, $l_2 = c_2 = p_2x_2$ denote the transition probability functions of routes in $B \to E$ respectively.

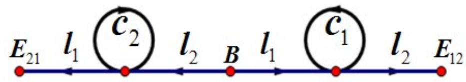

Figure 4: The TPFG of $G(p_1, p_2)$

Hence, by Lemma 1, we get the joint probability generating function of (ξ_1, ξ_2), that is the occurrence number of A_1 and A_2, as follows

$$G_{\xi_1,\xi_2}(x_1, x_2) = \frac{p_1p_2x_1x_2}{1-p_1x_1} + \frac{p_1p_2x_1x_2}{1-p_2x_2}. \quad (14)$$

Let $x_1 = x_2 = x$ in (14), by Lemma 3, the joint generating function of $\xi = \xi_1 + \xi_2$ is given by

$$G_\xi(x) = \frac{p_1p_2x^2}{1-p_1x} + \frac{p_1p_2x^2}{1-p_2x}. \quad (15)$$

Using Lemma 4, we can derive the mean and variance of ξ from formula (15).

Furthermore, expanding $G_\xi(x)$ into power series, we obtain

$$G_\xi(x) = \sum_{n=2}^{\infty} p_1p_2(p_1^{n-2} + p_2^{n-2})x^n. \quad (16)$$

Thus, from (16), we have the probability distribution of ξ as follows

$$P(\xi = n) = p_1p_2(p_1^{n-2} + p_2^{n-2}), n = 2, 3, \cdots.$$

Theorem 22 has been proved. □

Theorem 23 *Let $\xi = \xi_1 + \xi_2 + \xi_3$ be the generalized geometric random variable with parameter vector (p_1, p_2, p_3), then*

$$E(\xi) = \sum_{1\le i\le 3}\frac{1}{p_i} - \sum_{1\le i<j\le 3}\frac{1}{p_i+p_j} + 1, \quad (17)$$

$$Var(\xi) = \sum_{1\le i\le 3}\frac{1-p_i}{p_i^2} - 3\sum_{1\le i<j\le 3}\frac{1-(p_i+p_j)}{(p_i+p_j)^2}$$
$$-2\frac{1-(p_1p_2+p_1p_3+p_2p_3)}{(p_1+p_2)(p_1+p_3)(p_2+p_3)}, \quad (18)$$

$$E(\xi_k) = p_k \cdot E(\xi), k = 1, 2, 3, \quad (19)$$

$$P(\xi = n+3) = p_1p_2p_3\sum_{r=0}^{n}\sum_{1\le i<j\le 3}$$
$$(p_i^r + p_j^r)(p_i+p_j)^{n-r}, n = 0, 1, 2, \cdots. \quad (20)$$

Proof: In the TPFG of the trial process (see Figure 5), B is the beginning state, $E_{123} = \cdots = E_{321} = E$ is the ending state. There are $3! = 6$ parallel routes, with the transition probability functions: $l_k = c_k = p_k x_k$, $c_{ij} = p_i x_i + p_j x_j$, $k, i, j = 1, 2, 3, i \neq j$.

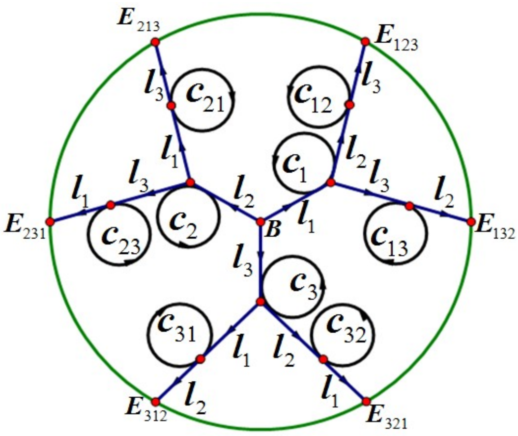

Figure 5: The TPFG of $G(p_1, p_2, p_3)$

Let g_{123} be the transition probability function of the route $B \to E_{123}$, then it follows from Lemma 2 that

$$g_{123} = \frac{p_1 p_2 p_3 x_1 x_2 x_3}{(1 - p_1 x_1)(1 - p_1 x_1 - p_2 x_2)}.$$

Similarly for $B \to E_{132}, \cdots, B \to E_{321}$, we have their transition probability functions

$$g_{132} = \frac{p_1 p_2 p_3 x_1 x_2 x_3}{(1 - p_1 x_1)(1 - p_1 x_1 - p_3 x_3)},$$

$$g_{213} = \frac{p_1 p_2 p_3 x_1 x_2 x_3}{(1 - p_2 x_2)(1 - p_2 x_2 - p_1 x_1)},$$

$$g_{231} = \frac{p_1 p_2 p_3 x_1 x_2 x_3}{(1 - p_2 x_2)(1 - p_2 x_2 - p_3 x_3)},$$

$$g_{312} = \frac{p_1 p_2 p_3 x_1 x_2 x_3}{(1 - p_3 x_3)(1 - p_3 x_3 - p_1 x_1)},$$

$$g_{321} = \frac{p_1 p_2 p_3 x_1 x_2 x_3}{(1 - p_3 x_3)(1 - p_3 x_3 - p_2 x_2)}.$$

By Lemma 1, we obtain the joint probability generating function of (ξ_1, ξ_2, ξ_3) as

$$G_{\xi_1,\xi_2,\xi_3}(x_1, x_2, x_3) = g_{123} + g_{132} + g_{213} + g_{231} + g_{312} + g_{321}. \quad (21)$$

Let $x_1 = x_2 = x_3 = x$ in (21), by Lemma 3, we have the probability generating function of $\xi = \xi_1 + \xi_2 + \xi_3$ as follows

$$G_\xi(x) = G_{\xi_1,\xi_2,\xi_3}(x, x, x),$$

and by Lemma 4, we can derive the mean and variance of ξ as equations (17) and (18).

Let $x_2 = x_3 = 1$ in (21), then we get the probability generating function of ξ_1

$$G_{\xi_1}(x_1) = G_{\xi_1,\xi_2,\xi_3}(x_1, 1, 1),$$

similarly,

$$G_{\xi_2}(x_2) = G_{\xi_1,\xi_2,\xi_3}(1, x_2, 1),$$

$$G_{\xi_3}(x_3) = G_{\xi_1,\xi_2,\xi_3}(1, 1, x_3).$$

Hence, by $E\xi_k = G'_{\xi_k}(1), k = 1, 2, 3$, we can derive the mean of ξ_k.

Similar to formula (16), we have

$$\begin{aligned} G_\xi(x) &= p_1 p_2 p_3 \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} p_1^r (p_1 + p_2)^{n-r} \right) x^{n+3} \\ &+ p_1 p_2 p_3 \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} p_1^r (p_1 + p_3)^{n-r} \right) x^{n+3} \\ &+ p_1 p_2 p_3 \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} p_2^r (p_2 + p_1)^{n-r} \right) x^{n+3} \\ &+ p_1 p_2 p_3 \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} p_2^r (p_2 + p_3)^{n-r} \right) x^{n+3} \\ &+ p_1 p_2 p_3 \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} p_3^r (p_3 + p_1)^{n-r} \right) x^{n+3} \\ &+ p_1 p_2 p_3 \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n} p_3^r (p_3 + p_2)^{n-r} \right) x^{n+3}. \end{aligned} \quad (22)$$

From formula (22), we get the probability distribution of ξ as follows

$$P(\xi = n + 3) = p_1 p_2 p_3 \sum_{r=0}^{n} \sum_{1 \le i < j \le 3} (p_i^r + p_j^r)(p_i + p_j)^{n-r}, n = 0, 1, 2, \cdots.$$

The proof is complete. □

Theorem 24 *If $\xi = \xi_1 + \xi_2 + \xi_3 + \xi_4$ is distributed as $G(p_1, p_2, p_3, p_4)$, then we have*

$$E\xi = \sum_{r=1}^{4} (-1)^{r-1} \sum_{1 \le l_1 < l_2 < \cdots < l_r \le 4} \frac{1}{p_{l_1} + p_{l_2} + \cdots + p_{l_r}}, \quad (23)$$

$$G''_\xi(1) = 2 \sum_{r=1}^{4} (-1)^{r-1} \sum_{1 \le l_1 < l_2 < \cdots < l_r \le 4} \frac{1 - (p_{l_1} + p_{l_2} + \cdots + p_{l_r})}{(p_{l_1} + p_{l_2} + \cdots + p_{l_r})^2}, \quad (24)$$

$$E\xi_k = p_k \cdot E\xi, k = 1, 2, 3, 4. \quad (25)$$

Proof: Following the operation of $G(p_1, p_2, p_3, p_4)$, we get the TPFG as shown in Figure 6.

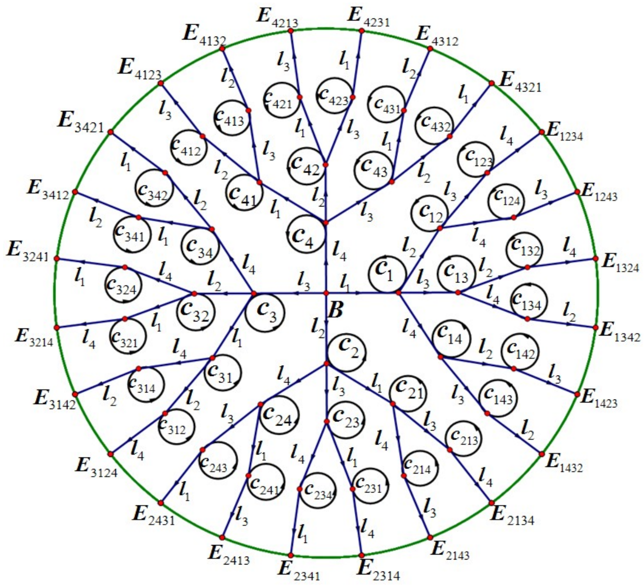

Figure 6: The TPFG of $G(p_1, p_2, p_3, p_4)$

Where B is the beginning state, $E_{1234} = \cdots = E_{4321} = E$ is the ending state. There are $4! = 24$ parallel routes, with the transition probability functions:

$$l_k = c_k = p_k x_k, k = 1, 2, 3, 4,$$

$$c_{ij} = p_i x_i + p_j x_j, i, j = 1, 2, 3, 4, i \neq j,$$

$$c_{ijk} = p_i x_i + p_j x_j + p_k x_k, i, j, k = 1, 2, 3, 4, i \neq j \neq k.$$

We put the 24 routes into 4 sets by l_1, l_2, l_3, l_4, for instance, $\{B \to E_{1234}, B \to E_{1243}, B \to E_{1324}, B \to E_{1342}, B \to E_{1423}, B \to E_{1432}\}$ is one of the 4 sets. We get its joint transition probability function as follows

$$\begin{aligned} & g_1(x_1, x_2, x_3, x_4) = \\ & \frac{p_1p_2p_3p_4x_1x_2x_3x_4}{(1-p_1x_1)(1-p_1x_1-p_2x_2)(1-p_1x_1-p_2x_2-p_3x_3)} \\ & + \frac{p_1p_2p_4p_3x_1x_2x_4x_3}{(1-p_1x_1)(1-p_1x_1-p_2x_2)(1-p_1x_1-p_2x_2-p_4x_4)} \\ & + \frac{p_1p_3p_2p_4x_1x_3x_2x_4}{(1-p_1x_1)(1-p_1x_1-p_3x_3)(1-p_1x_1-p_3x_3-p_2x_2)} \\ & + \frac{p_1p_3p_4p_2x_1x_3x_4x_2}{(1-p_1x_1)(1-p_1x_1-p_3x_3)(1-p_1x_1-p_3x_3-p_4x_4)} \\ & + \frac{p_1p_4p_2p_3x_1x_4x_2x_3}{(1-p_1x_1)(1-p_1x_1-p_4x_4)(1-p_1x_1-p_4x_4-p_2x_2)} \\ & + \frac{p_1p_4p_3p_2x_1x_4x_3x_2}{(1-p_1x_1)(1-p_1x_1-p_4x_4)(1-p_1x_1-p_4x_4-p_3x_3)}. \end{aligned}$$

In addition, we can get

$$g_2(x_1, x_2, x_3, x_4),$$

$$g_3(x_1, x_2, x_3, x_4),$$

$$g_4(x_1, x_2, x_3, x_4),$$

that is, the joint transition probability functions of the other 3 sets, where we omit their expressions similarly to $g_1(x_1, x_2, x_3, x_4)$.

Hence, by Lemma 1, we have the joint probability generating function of $(\xi_1, \xi_2, \xi_3, \xi_4)$

$$G_{\xi_1,\xi_2,\xi_3,\xi_4}(x_1, x_2, x_3, x_4) = \sum_{k=1}^{4} g_k(x_1, x_2, x_3, x_4).$$

By Lemma 4, we get the probability generating function of ξ in the following

$$G_\xi(x) = \sum_{k=1}^{4} g_k(x, x, x, x).$$

By computing the first and second derivative of $G_\xi(x)$ at $x = 1$, we get equations (23) and (24).

Note that

$$G_{\xi_1}(x_1) = G_{\xi_1,\xi_2,\xi_3,\xi_4}(x_1, 1, 1, 1), \cdots,$$
$$G_{\xi_4}(x_4) = G_{\xi_1,\xi_2,\xi_3,\xi_4}(1, 1, 1, x_4)$$

and

$$E\xi_k = G'_{\xi_k}(1), k = 1, 2, 3, 4,$$

we shall get (25). This completes the proof. □

When the parameter $n = 5$, there are $5!$ routes in the TPFG of $G(p_1, p_2, p_3, p_4, p_5)$ from beginning to ending. It's difficult to get the TPFG and the corresponding probability generating function. We find that the computational complexity increases rapidly in parameter n. In fact, even if $n = 4$ as stated above, the description of TPFG and the presentation for the probability generating function are very complicated. Nevertheless, the work for $n = 4$ is necessary, since it is hard to summarize $G(p_1, p_2, \cdots, p_n)$ only from $n = 2, 3$. By generalizing the results in Theorems 22, 23 and 24, we shall get

Theorem 25 *If the random variable $\xi = \sum_{j=1}^{n} \xi_j$ is distributed as $G(p_1, \cdots, p_n)$, then*

$$E\xi = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \le l_1 < l_2 < \cdots < l_r \le n} \frac{1}{p_{l_1} + p_{l_2} + \cdots + p_{l_r}}, \tag{26}$$

$$G''_\xi(1) = 2\sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \le l_1 < l_2 < \cdots < l_r \le n} \frac{1 - (p_{l_1} + p_{l_2} + \cdots + p_{l_r})}{(p_{l_1} + p_{l_2} + \cdots + p_{l_r})^2}, \tag{27}$$

$$E\xi_j = p_j \cdot E\xi, j = 1, 2, \cdots, n. \tag{28}$$

Remark 26 *1) There are $(2^n - 1)$ terms in the right-hand side of formula (26), each term is a mean of some random variable. All the terms can be divided into n classes: $\{\frac{1}{p_l}; 1 \le l \le n\}, \{\frac{1}{p_{l_1}+p_{l_2}}; 1 \le l_1 < l_2 \le n\}, \{\frac{1}{p_{l_1}+p_{l_2}+p_{l_3}}; 1 \le l_1 < l_2 < l_3 \le n\}, \cdots, \{\frac{1}{p_1+p_2+\cdots+p_n} = 1\}$.*

2) Following equations (26) and (27), we can get the variance of ξ. Especially, if ξ is distributed as $G(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$, then its variance is

$$\sum_{r=1}^{n} \frac{(-1)^{r-1}\binom{n}{r} n(2n-r)}{r^2} - \left[\sum_{r=1}^{n} \frac{(-1)^{r-1}\binom{n}{r} n}{r}\right]^2.$$

3) By combining equation (26) with the result presented in Gao's work [9], we obtain a combinatorial identity

$$\sum_{r=1}^{n} \frac{(-1)^{r-1}\binom{n}{r}}{r} = \sum_{r=1}^{n} \frac{1}{r}.$$

5 Conclusions

In order to discuss the probability properties of some success runs in Bernoulli trials, we introduce the transition probability flow graphs methods. In sections 2 and 3, following the TPFG methods we derive the generating functions of the geometric distribution of order k and the negative binomial distribution of order k. Then we get the means, variances and probability distributions, the exact probability distributions of them, furthermore, we show the relation of $G_k(p)$ and $NB_k(p)$. In section 4, we define the generalized geometric distribution $G(p_1, \cdots, p_n)$, discuss its mean, variance, probability distribution and some other characteristics by multivariate transition probability flow graphs theory. Moreover, we believe that TPFG methods should also be applied to other fields, for example, see [13], [20] and [21].

Acknowledgements: This work is supported by National Natural Science Foundation of China (No. 71271026) and the Fundamental Research Funds for the Central Universities (S11JB00400), during which this work was carried out. The authors are particularly indebted to the referees for reading the manuscript carefully and for giving valuable comments and suggestions.

References:

[1] S. Aki, K. Hirano, Distributions of numbers of failures and successes until the first consecutive k successes, *Annals of the Institute of Statistical Mathematics*, 46(1), 1994, pp. 193-202.

[2] M. J. Barry, A. J. L. Bello, The moment generating function of the geometric distribution of order k, *The Fibonacci Quarterly*, 31(2), 1993, pp. 178-180.

[3] C. S. Dai, Y. Y. Hou, The relationships between the negative hypergeometric distribution, the negative binomial distribution and the Poisson distribution, and their generalization, *Journal of Nanjing University mathematical Biquarterly*, 18(1), 2001, pp. 76-84.

[4] Y. L. Fan, The probability theoretical basis and application methods for TPFG(I), *Application of Statistics and Management*, 17(1), 1998, pp. 45-51.

[5] Y. L. Fan, The probability theoretical basis and application methods for TPFG(II), *Application of Statistics and Management*, 17(2), 1998, pp. 53-59.

[6] Y. L. Fan, The probability theoretical basis and application methods for TPFG(III), *Application of Statistics and Management*, 17(3), 1998, pp. 55-60.

[7] W. Feller, *An Introduction to Probability Theory and Its Applications,* vol.1, (Wiley, New York, 3rd ed), 1968.

[8] J. C. Fu, M. V. Koutras, Distributions theory of runs:A Markov chain approach, *Journal of the American Statistical Association*, 89(427), 1994, pp. 1050-1058.

[9] H. X. Gao, P. Niu, Y. Xue, L. Su, Distribution of sum of geometric distributions with different parameters, *Studies in college mathematics*, 16(1), 2013, pp. 21-22.

[10] A. P. Godbole, Specific formulae for some success run distributions, *Statist. Probab. Lett.*, 10(2), 1990, pp. 119-124.

[11] Q. Han, Exact distribution theory of runs, *Chinese Journal of Applied Probability and Statistics*, 15(2), 1999, pp. 199-212.

[12] T. Koyama, Switching characteristics under MIL-STD-105D, *Technometrics*, 21(1), 1979, p-p. 9-19.

[13] M. S. Khan, G. R. Pasha, A. H. Pasha, Theoretical analysis of inverse weibull distribution, *WSEAS Transactions on Mathematics*, 7(2), 2008, pp. 30-38.

[14] S. J. Mason, Feedback theory: Some properties of signal flow graphs, *Proceedings of the IRE.*, 41(9), 1953, pp. 1144-1156.

[15] M. Muselli, Simple expressions for success run distributions in Bernoulli trials, *Statist. Probab. Lett.*, 31(2), 1996, pp. 121-128.

[16] A. N. Philippou, C. Georghiou, G. N. Philippou, A generalized geometric distribution and some of its properties, *Statist. Probab. Lett.*, 1(4), 1983, pp. 171-175.

[17] A. N. Philippou, F. S. Makri, Success, runs and longest runs, *Statist. Probab. Lett.*, 4(4), 1986, pp. 211-215.

[18] S. J. Schwager, Run probabilities in sequences of Markov-dependent trials, *Journal of the American Statistical Association*, 78(381), 1983, pp. 168-175.

[19] J. G. Shao, Numerical characteristics for the number of independent trials, *Journal of Beijing Jiaotong University*, 38(6), 2014, pp. 106-111.

[20] J. G. Shao, J. Wang, Analysis of chain reaction between two stock indices fluctuations by statistical physics systems, *WSEAS Transactions on Mathematics*, 9(11), 2010, pp. 830-839.

[21] J. Wang, C. N. Wei, The asymptotical behavior of probability measures for the fluctuations of stochastic models, *WSEAS Transactions on Mathematics*, 7(5), 2008, pp. 273-282.