Ann. Inst. Statist. Math.
36 (1984), Part A, 431-440

ON DISCRETE DISTRIBUTIONS OF ORDER k

SIGEO AKI, HISATAKA KUBOKI AND KATUOMI HIRANO

(Received Dec. 14, 1983; revised Apr. 18, 1984)

Summary

This paper gives some results on calculation of probabilities and moments of the discrete distributions of order k. Further, a new distribution of order k, which is called the logarithmic series distribution of order k, is investigated. Finally, we discuss the meaning of the *order* of the distributions.

1. Introduction

Philippou, Georghiou and Philippou [6] introduced some distributions of order k such as the geometric, the negative binomial and the Poisson distribution of order k. They are defined as follows:

(i) The geometric distribution of order k

A random variable X is said to have the geometric distribution of order k with parameter p, to be denoted by $G_k(p)$, if

$$\mathrm{P}(X=x)=\sum_{x_1,\cdots,x_k}\binom{x_1+\cdots+x_k}{x_1,\ \cdots,\ x_k}p^x\left(\frac{q}{p}\right)^{x_1+\cdots+x_k},\qquad x\geqq k,$$

where the summation is over all nonnegative integers $x_1,\cdots,x_k$ such that

$$x_1+2x_2+\cdots+kx_k=x-k,\qquad\text{and}\qquad q=1-p.$$

(ii) The negative binomial distribution of order k

A random variable X is said to have the negative binomial distribution of order k with parameters p and r, to be denoted by $NB_k(r, p)$, if

$$\mathrm{P}(X=x)=\sum_{x_1,\cdots,x_k}\binom{x_1+\cdots+x_k+r-1}{x_1,\ \cdots,\ x_k,\ r-1}p^x\left(\frac{q}{p}\right)^{x_1+\cdots+x_k},\qquad x\geqq kr,$$

Key words: Discrete distributions, generalized geometric distribution, generalized negative binomial distribution, generalized Poisson distribution, generalized logarithmic series distribution.

where the summation is over all nonnegative integers $x_1, \cdots, x_k$ such that

$$x_1+2x_2+\cdots+kx_k=x-kr .$$

This is the distribution of sum of r i.i.d. random variables distributed as $G_k(p)$.

(iii) The Poisson distribution of order k

A random variable X is said to have the Poisson distribution of order k with parameter λ, to be denoted by $P_k(\lambda)$, if

$$\mathrm{P}(X=x)=\sum_{x_1,\cdots,x_k} e^{-k\lambda}\frac{\lambda^{x_1+\cdots+x_k}}{x_1!\cdots x_k!}, \qquad x=0, 1, \cdots ,$$

where the summation is over all nonnegative integers $x_1, \cdots, x_k$ such that

$$x_1+2x_2+\cdots+kx_k=x .$$

Philippou, Georghiou and Philippou [6] proposed that the number of trials until the occurrence of the kth consecutive success in independent trials with success probability p is distributed as $G_k(p)$. They also derived its mean and variance and the probability generating function from the definition of $G_k(p)$. Of course, they coincide with those of the number of trials until the occurrence of the kth consecutive success in independent trials (see Feller [2]). Further, Philippou and Muwafi [7] discussed the relationship between the geometric distribution of order k and the Fibonacci sequence of order k.

It is sometimes convenient to consider the shifted $G_k(p)$ and the shifted $NB_k(r, p)$ distributions so that the support of each distribution should become the set $\{0, 1, 2, \cdots\}$. We shall denote by $\overline{G}_k(p)$ and $\overline{NB}_k(r, p)$, respectively, the shifted distributions defined above.

In Section 2, we shall give some useful formulas for calculating their probabilities and moments. In Section 3, we consider a new type of distribution of order k, which is called the logarithmic series distribution of order k. In the last section, we shall discuss what the meaning of the *order* of each distribution of order k is. It is explained to be the degree of the polynomial which is the probability generating function of the generalizer in the 'α-generalization', which determines each distribution of order k.

2. Calculation of probabilities and moments

It is not easy to calculate the probabilities of the distributions of order k from the definitions. For the geometric distribution of order

k, however, the following proposition is useful.

PROPOSITION 2.1. Let X be a random variable distributed as $G_k(p)$. Then we have

$$(2.1)\qquad P(X=x)=\begin{cases} 0 & \text{for } x<k\,, \\ p^k & \text{for } x=k\,, \\ qp^k\left(1-\sum_{i=0}^{x-k-1} P(X=i)\right) & \text{otherwise}\,. \end{cases}$$

PROOF. Let E_x be the event that the first run of successes of length k occurs at the xth trial in independent trials with success probability p. And let F_x be the event that a run of successes of length k occurs at the xth trial and $(x-k)$th outcome is failure when $x-k>0$. It is easily seen that

$$P(F_x)=\begin{cases} p^k q & \text{for } x\geqq k+1\,, \\ p^k & \text{for } x=k\,, \\ 0 & \text{for } x<k\,. \end{cases}$$

Then, (2.1) follows by considering that

$$P(E_x)=P(F_x)-\sum_{i=0}^{x-k-1} P(E_i\cap F_x) \qquad \text{for } x\geqq k+1\,.$$

The formula (2.1) can be available for calculating the probability of the negative binomial distribution of order k, since it is the distribution of sum of r independent identically distributed random variables from the geometric distribution of order k.

As for the Poisson distribution of order k, we shall treat it as a special case of the following *generalized* Poisson distribution.

Adelson [1] investigated the *stuttering* Poisson distribution. We introduce it in a slightly extended form.

Let $\{X_i\}$ be a sequence of independent random variables each one of which has the distribution determined by

$$P(X_i=ik)=\exp(-\lambda_i)\frac{\lambda_i^k}{k!}\,, \qquad (k=0, 1, 2, \cdots)\,.$$

Assuming that $\lambda=\sum_{i=1}^{\infty}\lambda_i<\infty$, we shall call the distribution of $X=\sum_{i=1}^{\infty}X_i$ the *generalized* Poisson distribution with parameter $\boldsymbol{\lambda}=(\lambda_1, \lambda_2, \cdots)$. The probability of X is represented as

$$P(X=i)=\sum_{k_1,\cdots,k_i}\exp(-\lambda)\frac{\lambda_1^{k_1}\lambda_2^{k_2}\cdots\lambda_i^{k_i}}{k_1!\,k_2!\cdots k_i!}\,,$$

where the summation is over all nonnegative integers $k_1, k_2, \cdots, k_i$ such that

$$k_1+2k_2+\cdots+ik_i=i .$$

If we put $\lambda_1=\lambda_2=\cdots=\lambda_k$ and $0=\lambda_{k+1}=\lambda_{k+2}=\cdots$, then we get the Poisson distribution of order k. The *stuttering* Poisson distribution is also a special case of the *generalized* Poisson distribution.

We note that the *generalized* Poisson distribution can be reduced to a distribution which is obtained by the *usual generalization* of the Poisson distribution by a transform of the parameter.

Before showing that, we shall remark on the *usual generalization* of distributions, especially on that of the Poisson distribution.

Let $\phi_1(t)$ and $\phi_2(t)$ be the probability generating functions of two distributions F_1 and F_2, respectively. Then the distribution whose probability generating function is $\phi_1(\phi_2(t))$ is called a generalized F_1 distribution (generalized) by the generalizer F_2. When F_1 is the Poisson distribution, the generalized F_1 distribution has remarkable properties about its cumulants and the value of probability at k, which is denoted by $P(k)$. Since the probability generating function is $\phi(t)=\exp(\lambda(\phi_2(t)-1))$, the cumulant generating function is $\lambda\phi_2(e^t)-\lambda$. By differentiating it k times, we have

$$\kappa_k=\lambda\mu'_k(F_2) , \tag{2.2}$$

where κ_k is the kth cumulant and $\mu'_k(F_2)$ is the kth moment of the distribution F_2. Moreover, $\phi(t)$ satisfies the recurrence relation

$$\begin{aligned}\frac{1}{(k+1)!}(d/dt)^{k+1}\phi(t)&=\frac{\lambda}{(k+1)!}(d/dt)^k(\phi(t)\phi'_2(t))\\&=\frac{\lambda}{k+1}\sum_{j=0}^{k}(j+1)\left(\frac{\phi^{(k-j)}(t)}{(k-j)!}\right)\left(\frac{\phi_2^{(j+1)}(t)}{(j+1)!}\right) .\end{aligned}$$

From this, the values of $\mathrm{P}(k)$ are calculated by the following formula

$$\mathrm{P}(k+1)=\frac{\lambda}{k+1}\sum_{j=0}^{k}(j+1)\,\mathrm{P}(k-j)\,\mathrm{P}_2(j+1) , \tag{2.3}$$

where $\mathrm{P}_2(j)$ is the probability function of F_2.

Suppose we are given a *generalized* Poisson distribution whose probability generating function is written as

$$\phi(t)=\exp\left(-\sum_{i=1}^{\infty}\lambda_i+\sum_{i=1}^{\infty}\lambda_i t^i\right) .$$

Since $\lambda=\sum_{i=1}^{\infty}\lambda_i<\infty$, we put $\mu_i=\lambda_i/\lambda$. Then we have

$$\phi(t)=\exp\left(\lambda\left(\sum_{i=1}^{\infty}\mu_i t^i-1\right)\right) .$$

This means that the distribution with probability generating function $\phi(t)$ is the generalization of the Poisson distribution with parameter λ and the probability generating function of the generalizer is given by $\phi_2(t)=\sum_{i=1}^{\infty}\mu_i t^i$. Therefore, we can calculate the probability and the cumulants of the *generalized* Poisson distribution by using (2.2) and (2.3), respectively. When the generalizer has the jth moment, the kth cumulant of the *generalized* Poisson distribution is

$$\kappa_k=\sum_{i=1}^{\infty} i^k\lambda_i , \qquad k=1, 2,\cdots, j .$$

Recently, Philippou [5] discussed another distribution of order k, which he called "the compound distribution of order k".

3. Logarithmic series distribution of order k

The next proposition was obtained by Philippou, Georghiou and Philippou [6].

PROPOSITION 3.1. Let Y be a random variable distributed as $NB_k(r, p)$ and assume that $p\to 1$ and $r(1-p)\to\lambda$ (>0) as $r\to\infty$. Then

$$\mathrm{P}(Y-kr=y)\to\sum_{y_1,\cdots,y_k}\exp(-k\lambda)\frac{\lambda^{y_1+\cdots+y_k}}{y_1!\cdots y_k!} ,$$

where the summation is over all nonnegative integers $y_1,\cdots, y_k$ such that

$$y_1+2y_2+\cdots+ky_k=y .$$

This means that the Poisson distribution of order k is a limiting form of the negative binomial distribution of order k. When $k=1$, the relation between the two distributions is well known.

Now we shall show that a distribution exists which is also a limiting form of the negative binomial distribution of order k.

It is naturally seen that the negative binomial distribution of order k is defined for any positive real number r by

$$\mathrm{P}(X=x)=\sum_{x_1,\cdots,x_k}\binom{x_1+\cdots+x_k+r-1}{x_1,\ \cdots,\ x_k,\ r-1}p^{x+kr-[kr]}\left(\frac{q}{p}\right)^{x_1+\cdots+x_k} , \qquad x\geqq[kr] ,$$

where the summation is over all nonnegative integers $x_1,\cdots, x_k$ such that

$$x_1+2x_2+\cdots+kx_k=x-[kr] .$$

PROPOSITION 3.2. Let X be a random variable distributed as $NB_k(r, p)$ and assume that $r\to 0$. Then

$$P(X=x\mid X\geqq[kr]+1)\to\sum_{x_1,\cdots,x_k}\frac{(x_1+\cdots+x_k-1)!}{x_1!\cdots x_k!}\frac{1}{-k\log p}p^x\left(\frac{q}{p}\right)^{x_1+\cdots+x_k},$$

where the summation is over all nonnegative integers $x_1,\cdots,x_k$ such that

$$x_1+2x_2+\cdots+kx_k=x\ .$$

PROOF. Noting that

$$P(X=[kr])=p^{kr}\ ,$$

we have

$$\begin{aligned}
&P(X=x\mid X\geqq[kr]+1)\\
&\quad=\frac{P(X=x,\ X\geqq[kr]+1)}{1-P(X=[kr])}\\
&\quad=\sum_{x_1+2x_2+\cdots+kx_k=x-[kr]}\binom{x_1+\cdots+x_k+r-1}{x_1,\ \cdots,\ x_k,\ r-1}p^{x+kr-[kr]}\left(\frac{q}{p}\right)^{x_1+\cdots x_k}\Big/(1-p^{kr})\\
&\quad=\sum_{x_1+2x_2+\cdots+kx_k=x-[kr]}\frac{(x_1+\cdots+x_k+r-1)\cdots(r+2)(r+1)}{x_1!\cdots x_k!}\\
&\qquad\times\frac{rp^{kr}}{1-p^{kr}}p^{x_1+\cdots+kx_k}\left(\frac{q}{p}\right)^{x_1+\cdots+x_k}.
\end{aligned}$$

Notice that $\lim_{r\to 0}\frac{rp^{kr}}{1-p^{kr}}=\frac{-1}{k\log p}$. Then, it can be seen that $P(X=x\mid X\geqq[kr]+1)$ converges to

$$\sum_{x_1+2x_2+\cdots+kx_k=x}\frac{(x_1+\cdots+x_k-1)!}{(-k\log p)x_1!\cdots x_k!}p^x\left(\frac{q}{p}\right)^{x_1+\cdots+x_k},\qquad x=1,2,3,\cdots.$$

It is clear that the limit form of Proposition 3.2 is a probability distribution on the set $\{1, 2,\cdots\}$. Hence, we shall call it the logarithmic series distribution of order k. And we denote it by $LS_k(p)$. When $k=1$, the corresponding relation between the negative binomial distribution and the logarithmic series distribution is obtained by Fisher, Corbet and Williams [3].

The probability generating function and the ith moments about zero ($i=1, 2$) are respectively given by

$$\phi_{LS}(t)=\alpha(p)\log\left(\frac{1-pt}{1-t+qp^kt^{k+1}}\right),$$

$$\mu_1'=\alpha(p)(1-p^k-kqp^k)/(qp^k)$$

and

$$\mu_2' = \alpha(p)(1-p^{2k+1}-(2k+1)qp^k)/(q^2p^{2k}) \,, \qquad \text{where } \alpha(p) = -1/(k \log p) \,.$$

The next proposition is useful for calculation of the logarithmic series distribution of order k.

PROPOSITION 3.3. The values of probabilities of the logarithmic series distribution of order k at n ($n=1, 2, \cdots$), which are denoted by $P_{LS}(n)$, satisfy the following recursion formula

$$(3.1) \quad P_{LS}(n) = \alpha(p)[P_{\bar{G}}(n)/P_{\bar{G}}(0)] - \frac{1}{n}\sum_{j=1}^{n-1} j[P_{\bar{G}}(n-j)/P_{\bar{G}}(0)]\, P_{LS}(j) \,,$$

where $P_{\bar{G}}(i)$ is the value of probability of $\bar{G}_k(p)$ at i.

PROOF. The probability generating function of $\bar{G}_k(p)$ is given as

$$\phi_{\bar{G}}(t) = p^k(1-pt)/(1-t+qp^kt^{k+1}) \,.$$

Then, it is easily seen that

$$\phi_{LS}(t) = \alpha(p)[\log \phi_{\bar{G}}(t) - k \log p] \,.$$

By differentiating both sides, we have

$$(3.2) \qquad \phi_{\bar{G}}(t)\phi_{LS}^{(1)}(t) = \alpha(p)\phi_{\bar{G}}^{(1)}(t) \,.$$

Putting $t=0$, we get

$$P_{LS}(1) = \alpha(p)[P_{\bar{G}}(1)/P_{\bar{G}}(0)] \,.$$

Hence, (3.1) holds for $n=1$. For $n>1$, we differentiate both sides of (3.2) $(n-1)$ times, we obtain

$$(3.4) \qquad \sum_{j=0}^{n-1} \binom{n-1}{j} \phi_{\bar{G}}^{(n-j-1)}(t)\phi_{LS}^{(j+1)}(t) = \alpha(p)\phi_{\bar{G}}^{(n)}(t) \,.$$

If we set $t=0$ in (3.4), it holds that

$$\frac{1}{n}\sum_{j=0}^{n-2} (j+1)\, P_{\bar{G}}(n-j-1)\, P_{LS}(j+1) + P_{\bar{G}}(0)\, P_{LS}(n) = \alpha(p)\, P_{\bar{G}}(n) \,.$$

This implies (3.1).

4. Meaning of the "*order*"

In this section, we intend to give a systematic treatment of the *order* of the distributions by which we can explain the *order* of the distributions.

In the previous sections we introduced four distributions of *order* k, that is, the geometric, the negative binomial, the Poisson and the logarithmic series distribution of order k. The *order* of $G_k(p)$ is indeed explained as the length of the run of successes in independent trials with success probability p in the sense of Section 1. For the *order* of $NB_k(r, p)$, we can explain it in the same way. But we can hardly extend the explanation to the *order* of $P_k(\lambda)$, because it is only a limiting form of $NB_k(r, p)$ and we think it is not natural that $P_k(\lambda)$ is based on independent trials. We can not explain also the *order* of the logarithmic series distribution of order k based on independent trials because of the same reason.

Now, we explain what the meaning of the *order* of each distribution is, in other words, how we can get the distribution of order k from each distribution (of order 1) and whether the method is the same one for each distribution.

DEFINITION 4.1. Let $\psi_1(t)$ and $\psi_2(t)$ be the probability generating functions of two distributions F_1 and F_2, respectively. Let α be a positive real number which satisfies $\psi_1(\alpha)<\infty$. Then the distribution F whose probability generating function is equal to $\psi(t)=\psi_1(\alpha\psi_2(t))/\psi_1(\alpha)$ is called the α-generalized F_1 distribution by the generalizer F_2.

When $\alpha=1$, the α-generalization is the same as the usual generalization which was stated in Section 2. When F_1 is the Poisson distribution, the α-generalization is reduced to the usual generalization for any positive number α.

From the definition we can derive the moments of the α-generalized F_1 distribution by F_2.

PROPOSITION 4.1. If F_2 has its nth moment and $\psi_1^{(i)}(\alpha)$ $(=(d/dt)^i \psi_1(t)|_{t=\alpha})$ $(i=0, 1, \cdots, n)$ exist, then the nth moment of F is given as

$$\frac{1}{\psi_1(\alpha)}\sum_{r=0}^{n}\frac{\alpha^r\psi_1^{(r)}(\alpha)}{r!}\left[\sum_{s=0}^{r}\binom{s}{r}(-1)^{r-s}\sum_{i_1+\cdots+i_s=n}\frac{n!}{i_1!\cdots i_s!}\mu'_{i_1}\cdots\mu'_{i_s}\right],$$

where μ'_i is the ith moment of F_2 about zero.

PROOF. The moment generating function of F is written as

$$\phi(t)=\psi_1(\alpha\psi_2(e^t))/\psi_1(\alpha).$$

By differentiating both sides n times, we have

$$(d/dt)^n\phi(t)=\frac{1}{\psi_1(\alpha)}\sum_{r=0}^{n}\frac{\psi_1^{(r)}(\alpha\psi_2(e^t))}{r!}\left[\sum_{s=0}^{r}\binom{r}{s}(-\alpha\psi_2(e^t))^{r-s}(d/dt)^n(\alpha\psi_2(e^t))^s\right].$$

Thus,

$$(d/dt)^n\phi(t)|_{t=0}=\frac{1}{\psi_1(\alpha)}\sum_{r=0}^{n}\frac{\alpha^r\psi_1^{(r)}(\alpha)}{r!}\left[\sum_{s=0}^{r}\binom{r}{s}(-1)^{r-s}(d/dt)^n((\psi_2(e^t))^s|_{t=0}\right].$$

Note that $(\psi_2(e^t))^s$ is the moment generating function of the sum of s i.i.d. random variables distributed as F_2. Then it can be seen that

$$(d/dt)^n(\psi_2(e^t))^s|_{t=0}=\sum_{i_1+\cdots+i_s=n}\frac{n!}{i_1!\cdots i_s!}\mu'_{i_1}\cdots\mu'_{i_s}.$$

Then, the result is obvious.

Now, we propose a method for getting the distribution of order k from each distribution (of order 1). The probability generating functions of $\bar{G}_k(p)$, $\overline{NB}_k(r, p)$, $P_k(\lambda)$ and $LS_k(p)$ are written respectively as

(4.1) $$\psi_{\bar{G}}(t;k)=\frac{p^k(1-pt)}{1-t+qp^kt^{k+1}},$$

(4.2) $$\psi_{\overline{NB}}(t;k)=\left(\frac{p^k(1-pt)}{1-t+qp^kt^{k+1}}\right)^r,$$

(4.3) $$\psi_P(t;k)=\exp\left(-k\lambda+\lambda\sum_{i=1}^{k}t^i\right)$$

and

(4.4) $$\psi_{LS}(t;k)=\frac{-1}{k\log p}\log\left(\frac{1-pt}{1-t+qp^kt^{k+1}}\right).$$

Putting $\alpha_0=(1-p^k)/(1-p)$, we define

$$\psi_2(t;p)=\frac{1}{\alpha_0}(t+pt^2+\cdots+p^{k-1}t^k)=\frac{t}{\alpha_0}\cdot\frac{1-(pt)^k}{1-pt}.$$

It is clear that $\psi_2(t;p)$ is a probability generating function. Hence, we denote by $F_2(p)$ the distribution determined by the probability generating function $\psi_2(t;p)$. Then, the next proposition will be easily checked.

PROPOSITION 4.2. The next two statements hold concerning the construction of the distributions of order k.
(i) $\bar{G}_k(p)$, $\overline{NB}_k(r, p)$ and $LS_k(p)$ are obtained by the α_0-generalization of $\bar{G}_1(p)$, of $\overline{NB}_1(r, p)$ and of $LS_1(p)$, respectively, by the common generalizer $F_2(p)$.
(ii) $P_k(\lambda)$ is obtained by the α_0-generalization of $P_1(\lambda)$ by the generalizer $F_2(1)$.

From Proposition 4.2 we can explain that the *order* of each distribution is the degree of the polynomial which is the probability generating function of the generalizer in the sense of the proposition. And we shall note that the second statement of Proposition 4.2 is consistent with Proposition 3.1, since $P_k(\lambda)$ is a limiting form of $NB_k(r, p)$ as p goes to one.

Finally, we call attention to the shapes of the distributions of order k. They are very interesting and notable. Various graphs of discrete distributions including the distributions of order k are given by Hirano, Kuboki, Aki and Kuribayashi [4].

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Adelson, R. M. (1966). Compound Poisson distributions, *Operational Research Quarterly*, **17**, 73-75.

[2] Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. I (3rd ed.), John Wiley and Sons, New York.

[3] Fisher, R. A., Corbet, A. S. and Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population, *Journal of Animal Ecology*, **12**, 42-57.

[4] Hirano, K., Kuboki, H., Aki, S. and Kuribayashi, A. (1984). Figures of probability density functions in statistics II—discrete univariate case—, *Computer Science Monographs*, No. **20**, The Institute of Statistical Mathematics, Tokyo.

[5] Philippou, A. N. (1984). The Poisson and compound Poisson distribution of order k and some of their properties, *Zapiski Nauchnyka Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im*, V. A. Steklova AN SSSR, **130**, 175-180.

[6] Philippou, A. N., Georghiou, C. and Philippou, G. N. (1983). A generalized geometric distribution and some of its properties, *Statistics and Probability Letters*, **1**, 171-175.

[7] Philippou, A. N. and Muwafi, A. A. (1982). Waiting for the kth consecutive success and the Fibonacci sequence of order k, *The Fibonacci Quarterly*, **20**, 28-32.