Ann. Inst. Statist. Math.
38 (1986), Part A, 223-231

ON GENERALIZED BINOMIAL AND MULTINOMIAL DISTRIBUTIONS AND THEIR RELATION TO GENERALIZED POISSON DISTRIBUTIONS

JOHN PANARETOS AND EVDOKIA XEKALAKI

(Received Aug. 6, 1984; revised June 21, 1985)

1. Introduction

The binomial and multinomial distributions have found applications in many scientific fields, mainly because of their simplicity. Many practical problems however cannot be adequately described by the simple binomial or multinomial model. In such cases more elaborate binomial-type or multinomial-type models are needed. In this context many extensions and generalizations of the binomial and multinomial distributions have been defined and their properties have been studied.

The present paper describes such situations in the form of some urn schemes which lead to distributions more general than the usual binomial or multinomial distributions (Sections 2 and 4). In Section 3 properties relating the binomial-type distribution of Section 2 to generalized Poisson distributions are stated and proved. Finally, the obtained results are extended to multinomial-type distributions in Section 4.

2. The cluster binomial model and its probability distribution

Let us consider the following game of chance: Suppose that an urn contains balls. Each ball bears a number from 0 to k where k is

Key words: Cluster binomial distribution, cluster multinomial distribution, generalized Poisson distribution.

a fixed positive integer. The game consists of the player drawing n balls, with replacement, from the urn. The sum of the numbers on the balls drawn determines the amount of money the player wins, or looses at the end of the game. (The player, for example, can win, or loose, as many units as the difference between the above mentioned total and a number L determined by agreement beforehand).

Let X denote the sum of the numbers shown on the balls drawn. To study this situation we need to know the probability distribution of X.

Let p_i denote the probability that a ball bearing the number i will be drawn ($i=1, 2, \cdots, k$) with $\sum_{i=1}^{k} p_i \equiv p$. Then, $q=1-p$ is the probability that a ball bearing a zero will be drawn. Therefore, the random variable X takes the value r if, of the n balls drawn, r_1 bear the number 1, r_2 the number 2 and so on, r_k bear the number k so that $\sum_{i=1}^{k} ir_i = r$ and each of the remaining $n-\sum_{i=1}^{k} r_i$ balls bear the number zero. Consequently, the probability that the summation of the numbers on the balls drawn is r ($r=0, 1, 2, \cdots, nk$) is

$$\text{(2.1)} \qquad \mathrm{P}(X=r) = \sum_{\sum ir_i=r} \binom{n}{r_1, r_2, \cdots, r_k, n-\sum r_i} \left(\prod_{i=1}^{k} p_i^{r_i}\right) q^{n-\sum r_i} .$$

PROPOSITION 2.1. *The function defined by* (2.1) *is a proper probability distribution.*

PROOF. Using the transformation $r_i \to r_i$ and $r - \sum_{i=1}^{k} (i-1) r_i \to r$ we find

$$\begin{aligned} &\sum_{r=0}^{nk} \sum_{\sum ir_i=r} \binom{n}{r_1, r_2, \cdots, r_k, n-\sum r_i} \left(\prod_{i=1}^{k} p_i^{r_i}\right) q^{n-\sum r_i} \\ &= \sum_{r=0}^{n} \sum_{\sum r_i=r} \binom{n}{r_1, r_2, \cdots, r_k, n-\sum r_i} \left(\prod_{i=1}^{k} p_i^{r_i}\right) q^{n-\sum r_i} \\ &= \sum_{r=0}^{n} \binom{n}{r} q^{n-r} \sum_{\sum r_i=r} \binom{r}{r_1, r_2, \cdots, r_k} \left(\prod_{i=1}^{k} p_i^{r_i}\right) = 1 . \end{aligned}$$

Note. One can easily observe that for $k=1$ the distribution defined by (2.1) reduces to the ordinary binomial distribution.

DEFINITION 2.1. Let n, k be fixed positive integers and let p_i, $i=1, 2, \cdots, k$ be fixed numbers $0<p_i<1$. Set $p=\sum_{i=1}^{k} p_i$ and $q=1-p$. Consider a random variable X taking non-negative, integer values. The probability distribution defined by (2.1) is called cluster binomial distribution with parameters $n, k, p_1, p_2, \cdots, p_k$ and is denoted by $b_k(n; p_1, p_2, \cdots, p_k)$.

In the remaining of this section we derive $G(s)$, the probability generating function (p.g.f.), the mean and the variance of the distribution defined by (2.1).

PROPOSITION 2.2. *Let X follow the cluster binomial distribution with probabilities given by* (2.1). *Then*

$$\text{(i)}\quad G(s)=\left(q+\sum_{i=1}^{k}p_is^i\right)^n, \qquad |s|\leqq 1 \tag{2.2}$$

$$\text{(ii)}\quad \mathrm{E}(X)=n\sum_{i=1}^{k}ip_i \tag{2.3}$$

$$\text{(iii)}\quad V(X)=n\left(\sum_{i=1}^{k}i^2p_i-\left(\sum_{i=1}^{k}ip_i\right)^2\right). \tag{2.4}$$

The distribution defined in (2.1) was first arrived at by Steyn [5] as the limit of a generalization of the hypergeometric distribution.

It is evident from the form of the p.g.f. of the cluster binomial distribution as given by (2.2) that this distribution can be thought of as a binomial distribution generalized by a finite discrete distribution. More precisely, if N is a non-negative, integer-valued random variable following a binomial distribution with parameters n and p and $X_1, X_2, \cdots$ is a sequence of independently and identically distributed random variables, which are independent of N with probability distributions $\mathrm{P}(X_i=j)=p_i/p$, $j=1,2,\cdots,k$, $i=1,2,\cdots$, then the random variable $X=X_1+X_2+\cdots+X_N$ follows the cluster binomial distribution with parameters n, k and $p_1, p_2, \cdots, p_k$. This representation justifies the choice of the name cluster binomial for the distribution defined by (2.1) as it shows that this distribution can be viewed as the distribution of events that occur in clusters of size X_i, $1\leqq X_i\leqq k$. (Charalambides [2] provided an expression that permits the recursive evaluation of the probabilities of generalized binomial distributions in terms of the Bell polynomials).

Therefore, the cluster binomial distribution can have potential applications in many diverse fields of statistical analysis whenever there are multiple "groups" (clusters) of "items" and one is interested in the distribution of the total number of "items". It could be used, for example, in accident analysis as the distribution of the number of deaths resulting from n accidents with p_i representing the probability with which an accident will result in i deaths, $i=1,2,\cdots,k$ ($p_0=q$). On the other hand, the above representation of the cluster binomial distribution offers some other possibilities in the area of applications. As an illustration, consider a building with n elevators. Let k be the maximum number of people each of the elevators can take. Let p be the probability that an elevator is working at a given point in time and assume that this remains the same from elevator to elevator. It is not unreasonable to assume that the probability that the elevator i will be carrying X_i persons ($x_i=1, 2, \cdots, k$) is $1/k$. Then the total number X of people using the elevators at any given point in time will follow the cluster, binomial distribution with parameters n, k and $p_i=p/k$, $i=1, 2, \cdots, k$.

3. The relationship of the cluster binomial distribution to generalized Poisson distributions

The association of the ordinary binomial distribution to the ordinary Poisson distribution is well-known. The representation of the cluster binomial distribution, defined in Section 1, as a generalized binomial distribution makes one wonder whether there exists any association between the cluster binomial and the class of generalized Poisson distributions.

A random variable X is said to follow a generalized Poisson distribution if it can be written as the random sum of N independent, identically distributed random variables with N following a Poisson distribution. (Feller uses the term compound Poisson instead of generalized Poisson). It is well-known (see Feller [3], p. 291) that the p.g.f. of a generalized Poisson distribution $G(s)=\exp\{\lambda(g(s)-1)\}$, with λ the parameter of the Poisson distribution and $g(s)$ the p.g.f. of the common distribution of the X_j's, can always be written in the form

$$G(s)=\exp\left\{\sum_{i=1}^{k} \lambda_i(s^i-1)\right\} \tag{3.1}$$

where $\lambda_i=\lambda p_i$ and $p_i=\mathrm{P}(X_j=i)$, $i=1, 2, \cdots, k$; $k \in I^+ U\{+\infty\}$, $\sum_{i=1}^{k} \lambda_i < +\infty$.

Recently, the distribution in (3.1) was studied by Aki et al. [1] who refer to it as the generalized Poisson distribution with parameter $\boldsymbol{\lambda}=(\lambda_1, \lambda_2, \cdots)$.

In what follows some results are presented relating the cluster binomial distribution to the generalized Poisson distribution given by (3.1)

PROPOSITION 3.1. *Let X, Y be two non-negative, integer-valued random variables such that the conditional distribution of Y given that ($X=n$) is the cluster binomial distribution $b_k(n; p_1, p_2, \cdots, p_k)$ as in (2.1). Then, X follows a Poisson distribution with parameter λ if and only if Y follows a generalized Poisson distribution with parameters $\lambda p_1, \lambda p_2, \cdots, \lambda p_k$.*

PROOF. "Only if" part. From the assumptions it follows that

$$G_Y(s)=\sum_{x=0}^{\infty} G_{Y|(X=x)}(s)P(X=x)=\sum_{x=0}^{\infty}\left(q+\sum_{i=1}^{k} p_i s^i\right)^x e^{-\lambda}\frac{\lambda^x}{x!}=e^{\lambda\left(\sum_{i=1}^{k} p_i s^i-\sum_{i=1}^{k} p_i\right)} .$$

"If" part:

$$G_Y(s)=e^{\sum \lambda p_i(s^i-1)}=G_X\left(q+\sum_{i=1}^{k} p_i s^i\right) .$$

Making the transformation

$$q+\sum_{i=1}^{k} p_i s^i \rightarrow s \qquad \text{we get}$$

$$G_X(s)=e^{\lambda(\sum p_i s^i-\sum p_i)}=e^{\lambda(s-q-\sum p_i)}=e^{\lambda(s-1)} .$$

This completes the proof of the proposition.

PROPOSITION 3.2. *Let X be a non-negative, integer-valued r.v. that has the Poisson distribution with parameter $\lambda>0$. Let Y be another non-negative, integer-valued r.v. such that the conditional distribution of Y given $(X=n)$ is independent of λ. Then $Y|(X=n)$ follows the cluster binomial distribution with parameters $(n, p_1, p_2, \cdots, p_k)$ as in* (2.2) *if and only if Y follows a generalized Poisson distribution with parameter $(\lambda p_1, \lambda p_2, \cdots, \lambda p_k)$ and p.g.f. given by* (3.1).

PROOF. The "Only if" part holds by Proposition 3.1. For the "If" part observe that $B_n(s)$, the p.g.f. of the $b_k(n; p_1, p_2, \cdots, p_k)$, is one solution for $G_{Y|(X=n)}(s)$ in the equation

$$G_Y(s)=\sum_{n=0}^{\infty} G_{Y|(X=n)}(s)P(X=n) . \tag{3.2}$$

It will be shown that $B_n(s)$ is the only solution. Suppose that $B_n^*(s)$ is another solution of (3.2), independent of λ. Then, it is obvious from (3.2) that

$$\sum_{n=0}^{\infty}(B_n^*(s)-B_n(s))\frac{\lambda^n}{n!}=0$$

i.e. that $B_n^*(s)=B_n(s)$. Hence the result has been established.

For $k=1$, Proposition 3.2 is a special case of a result shown by Xekalaki and Panaretos [7].

PROPOSITION 3.3. *Let $X\sim b_k(n; p_1, p_2, \cdots, p_k)$ as in* (2.1). *Then as $n\rightarrow\infty$ and $p_i\rightarrow 0$ so that $np_i\rightarrow\lambda_i$, $i=1, 2, \cdots, k$ the $b_k(n; p_1, p_2, \cdots, p_k)$ tends to the generalized Poisson distribution with p.g.f. as in* (3.1).

PROOF. Let us denote by H the set of conditions $n\rightarrow\infty$, $p_i\rightarrow 0$ so that $np_i\rightarrow\lambda_i$. Using the inequalities $(t-1)/t\leqq\ln t\leqq t-1$ for $t>0$, we have

$$\frac{n(\sum p_i s^i - \sum p_i)}{1-\sum p_i + \sum p_i s^i} \leqq n \ln (q+\sum p_i s^i) \leqq n(\sum p_i s^i - \sum p_i) , \qquad q+\sum p_i s^i > 0$$

where $q=1-\sum p_i$. Taking the $\lim\limits_{H}$ yields

$$\sum \lambda_i s^i - \sum \lambda_i \leqq \lim_{H} n \ln (q+\sum p_i s^i) \leqq \sum \lambda_i s^i - \sum \lambda_i$$

and therefore,

$$\lim_{H} (q+\sum p_i s^i)^n = \exp\left(\sum_{i=1}^{k} \lambda_i (s^i - 1)\right) .$$

Hence the result has been established.

4. The cluster multinomial model and its probability distribution

As is well known, a natural extension of the urn scheme that gives rise to the usual binomial distribution leads to the usual multinomial distribution. In a like manner, one can extend the urn genesis scheme of the cluster binomial distribution to arrive at the definition of a multinomial type distribution, the cluster multinomial distribution.

The urn contains balls of m different colours. The balls of colour i are numbered from 0 to k_i, $i=1, 2, \cdots, m$. n balls are drawn at random with replacement. Let p_{ij} be the probability that a ball of colour i will bear number j, $j=0, 1, \cdots, k_i$; $i=1, 2, \cdots, m$ and let X_{ij} denote the number of balls of colour i in the sample that bear the number j. Then $X_i = \sum_{j=1}^{k_i} jX_{ij}$ represents the sum of the numbers shown on balls of colour i, $i=1, 2, \cdots, m$, and it can be shown that

(4.1) $$P(X_1=x_1, \cdots, X_m=x_m)$$
$$= \sum_{\sum_{j=1}^{k_i} j l_{ij} = x_i} \binom{n}{l_{11}, \cdots, l_{1k_1}, \cdots, l_{m1}, \cdots, l_{mk_m}} p_0^{\,n-\sum_{i=1}^{k_1+\cdots+k_m} l_i} \prod_{i=1}^{m} \prod_{j=1}^{k_1} p_{ij}^{l_{ij}} ,$$

$x_i=0, 1, \cdots, nk_i$, $i=1, 2, \cdots, m$.

One can easily verify that (4.1) defines a proper multivariate probability distribution which for $k_1=k_2=\cdots=k_m=1$ reduces to the ordinary multinomial distribution. In the sequel, we will refer to this distribution as the cluster multinomial distribution and denote it by $m_{\boldsymbol{k}}(n, \boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m)$ where $\boldsymbol{k}=(k_1, \cdots, k_m)$ and $\boldsymbol{p}_i=(p_{i1}, \cdots, p_{ik_i})$, $i=1, \cdots, m$.

PROPOSITION 4.1. *Let* $\boldsymbol{X} \equiv (X_1, X_2, \cdots, X_m)$ *follow the cluster multinomial distribution with parameters* n, $\boldsymbol{k}$ *and* $\boldsymbol{p}_i$, $i=1, 2, \cdots, m$. *Then the p.g.f. of* $\boldsymbol{X}$ *is given by*

$$G_X(t)=\left(p_0+\sum_{i=1}^{m}\sum_{j=1}^{k_i} p_{ij}t_i^j\right)^n . \tag{4.2}$$

PROOF. The proof follows exactly as in the univariate case.

Steyn [6] studied the asymptotic behaviour of a certain transformation of r.v.'s having a probability distribution generated by (4.2).

PROPOSITION 4.2. *Let* $\boldsymbol{X}=(X_1,\cdots,X_m)$ *be a random vector having the cluster multinomial distribution, i.e.* $\boldsymbol{X}\sim m_{\boldsymbol{k}}(n,\ \boldsymbol{p}_1,\cdots,\boldsymbol{p}_m)$. *Then*

(i) $X_i\sim b_{k_i}(n,\ p_{i1},p_{i2}\cdots,p_{ik_i}),\ i=1,2,\cdots,m.$

(ii) *In the case* $k_1=k_2=\cdots=k_m=k$

$\sum_{i=1}^{m} X_i\sim b_k\left(n,\ \sum_{i=1}^{m} p_{i1},\ \sum_{i=1}^{m} p_{i2},\cdots,\ \sum_{i=1}^{m} p_{ik}\right).$

PROOF. The results follow using (4.2) by setting $t_l=1,\ l=1,2,\cdots,m,\ l\neq i$ for (i) and $t_i=t,\ i=1,2,\cdots,m$ for (ii).

The form of the p.g.f. of the $m_{\boldsymbol{k}}(n,\ \boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m)$ implies that the cluster multinomial distribution can be thought of as the generalization of a multinomial distribution with parameters $n,p_1,p_2,\cdots,p_m$ where the i-th generalizer has p.g.f. $\sum_{j=1}^{k_i}(p_{ij}/p_i)t_i^j,\ i=1,2,\cdots,m.$

The results of Propositions 3.1, 3.2 and 3.3 can easily be extended to analogous results concerning the cluster multinomial distribution and its relationship to the generalized Poisson distribution as indicated by the propositions that follow.

PROPOSITION 4.3. *Let* $X,\ Y_1,Y_2,\cdots,Y_m$ *be non-negative, integer-valued r.v.'s such that the conditional distribution of* $(Y_1,Y_2,\cdots,Y_m)$ *given that* $(X=n)$ *is the cluster multinomial distribution* $m_{\boldsymbol{k}}(n;\ \boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m)$ *as in* (4.1). *Then X follows a Poisson distribution with parameter* $\lambda>0$, *if and only if* $Y_1,Y_2,\cdots,Y_m$ *are independent generalized Poisson r.v.'s with parameters* $\lambda\boldsymbol{p}_i=(\lambda p_{i1},\lambda p_{i2},\cdots,\lambda p_{ik}),\ i=1,2,\cdots,m$ *and p.g.f.'s.*

$$G_{Y_i}(s_i)=\exp\left\{\lambda\sum_{j=1}^{k} p_{ij}(s_i^j-1)\right\}, \qquad i=1,2,\cdots,m . \tag{4.3}$$

PROOF. "Only if" part. Let X be a Poisson (λ) r.v. Then

$$\begin{aligned}G_{Y_1,Y_2,\dots,Y_m}(s_1,\cdots,s_m)&=\sum_{n=0}^{\infty} G_{Y_1,\dots,Y_m|(X=n)}(s_1,\cdots,s_m)P(X=n)\\&=\sum_{n=0}^{\infty}\left(p_0+\sum_{i=1}^{m}\sum_{j=1}^{k} p_{ij}s_i^j\right)^n e^{-\lambda}\frac{\lambda^n}{n!}\\&=\prod_{i=1}^{m}\exp\left\{\lambda\sum_{j=1}^{k} p_{ij}(s_i^j-1)\right\}.\end{aligned}$$

"If" part. Suppose that $Y_1,\cdots,Y_m$ are independent generalized Poisson r.v.'s with p.g.f.'s given by (4.3). Then $Y=Y_1+Y_2+\cdots+Y_m$ follows a generalized Poisson distribution with parameter $\left(\lambda \sum_{i=1}^{m} p_{i1}, \lambda \sum_{i=1}^{m} p_{i2}, \cdots, \lambda \sum_{i=1}^{m} p_{ik}\right)$ and p.g.f.

$$G_Y(s)=\exp\left\{\lambda \sum_{j=1}^{k} \sum_{i=1}^{m} p_{ij}(s^j-1)\right\} .$$

But since $(Y_1, Y_2, \cdots, Y_m)|(X=n)$ has the $m_k(n, \boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m)$, it follows that $Y|(X=n)$ has the $b_k\left(n, \sum_{i=1}^{m} p_{i1}, \cdots, \sum_{i=1}^{m} p_{ik}\right)$.

Then, by Proposition 3.1, X follows a Poisson distribution with parameter λ. Hence the result has been established.

PROPOSITION 4.4. *Let X, $Y_1, Y_2, \cdots, Y_m$ be non-negative, integer-valued r.v.'s. Assume that X follows the Poisson distribution with parameter $\lambda>0$ and that the distribution of $(Y_1, Y_2, \cdots, Y_m)$ given that $(X=n)$ is independent of λ. Then $(Y_1, Y_2, \cdots, Y_m)|(X=n)$ follows the cluster multinomial distribution with parameters n, $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_m$ as in* (4.1) *if and only if $Y_1, Y_2, \cdots, Y_m$ are independent generalized Poisson r.v.'s with p.g.f.'s given by* (4.3).

PROOF. The result can be shown by an argument analogous to that used to prove Proposition 3.2.

By an analogous argument one can show the following result.

PROPOSITION 4.5. *Let $\boldsymbol{X} \sim m_k(n, \boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m)$ as in* (4.1). *Then as $n \to \infty$ and $\boldsymbol{p}_i \to \boldsymbol{0}$, $i=1, \cdots, m$ so that $n\boldsymbol{p}_i \to \boldsymbol{\lambda}_i < +\infty$, $i=1, 2, \cdots, m$, the $m_k(n, \boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m)$ tends to the joint distribution of m independent r.v.'s $Y_1, Y_2, \cdots, Y_m$ that follow generalized Poisson distributions with parameter vectors $\boldsymbol{\lambda}_i=(\lambda_{i1}, \lambda_{i2}, \cdots, \lambda_{ik_i})$, $i=1, 2, \cdots, m$.*

Acknowledgement

We would like to thank a referee for constructive comments on an earlier version of this paper.

UNIVERSITY OF PATRAS, UNIVERSITY OF CRETE (GREECE)
UNIVERSITY OF CRETE, ATHENS SCHOOL OF ECONOMICS (GREECE)

REFERENCES

[1] Aki, S., Kuboki, H. and Hirano, K. (1984). On discrete distributions of order k, *Ann. Inst. Statist. Math.*, **36**, 431–440.

[2] Charalambides, Ch. A. (1977). On the generalized discrete distributions and the Bell

polynomials, *Sankhyā*, B, **39**, 36-44.

[3] Feller, W. (1970). *An Introduction to Probability Theory and Its Applications*, (Revised printing of the third edition), Wiley, New York.

[4] Panaretos, J. and Xekalaki, E. (1984). Definition and properties of a generalized binomial distribution, *Technical Report No. 1* (June 1984), Department of Mathematics, University of Crete, Greece.

[5] Steyn, H. S. (1956). On the univariable series $F(t)=F(a;\ b_1, b_2, \cdots, b_k;\ c;\ t, t^2, \cdots, t^k)$ and its application in probability theory, *Proc. Kon. Ned. Akad. V. Wetensch.*, Ser. A, **59**, 190-197.

[6] Steyn, H. S. (1963). On approximations for the discrete distributions obtained from multiple events, *Proc. Kon. Ned. Akad. V. Wetensch.*, Series A, **66**, 85-96.

[7] Xekalaki, E. and Pararetos, J. (1983). Identifiability of compound Poisson distributions, *Scand. Actuarial J.*, **1983**, 39-45.