Multivariate Power Series Distributions and Neyman's Properties for Multinomials

J. K. Ghosh,* Bikas Kumar Sinha,† and Bimal Kumar Sinha*

* Indian Statistical Institute and University of Pittsburgh, and
† Calcutta University and Federal University of Bahia

Communicated by C. G. Khatri

A problem of J. Neyman (in Classical and Contagious Discrete Distributions (G. P. Patil, Ed.), 1965, pp. 4–14) regarding a characterization of positive and negative multinomial distributions is studied in this paper. Some properties of multivariate power series distributions in general which should be of independent interest are also derived.

1. Introduction and Summary

Neyman [4] obtained a number of interesting properties of the multivariate negative binomial distribution and raised the question whether the properties that he had found characterize this distribution. It is easy to see that the same properties also hold for the positive multinomial distribution. Moreover, if the properties hold for $X_1, ..., X_k$, they also hold for $Y_1, ..., Y_k$, where $Y_i = aX_i + b_i$, $i = 1, ..., k$. Hence, the interesting question is whether Neyman's properties characterize the two multinomials and distributions obtained from them by linear transformations of this type. But to make this question meaningful one first has to define what is meant by a family of distributions. This has to be done in such a way that the multinomials will occur as special cases. In the present context the following definition which is a slight variant of that of Sinha and Sinha [6] seems to be a reasonable one. For each n in some index set N, let $a_{i_1 \cdots i_k}^n \geqslant 0$ for $(i_1, ..., i_k) \in I^k = I \times \cdots \times I$, where I is the set of nonnegative integers. Let $\mathscr{H}_k = \{(\theta_1, ..., \theta_k): \theta_i > 0, i = 1, ..., k,$ and $\sum_{i_1, ..., i_k} a_{i_1 \cdots i_k}^n \theta_1^{i_1} \cdots \theta_k^{i_k} < \infty\}$. For each $\theta \in \mathscr{H}_k$, let

$$c(\theta) = \sum_{i_1, ..., i_k} a_{i_1 \cdots i_k}^n \theta_1^{i_1} \cdots \theta_k^{i_k}.$$

397

Consider $f(i_1,...,i_k \mid n, \boldsymbol{\theta}) = a^n_{i_1 \cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k}/c(\boldsymbol{\theta})$ and let $P(\cdot, n, \boldsymbol{\theta})$ be the corresponding probability measure. Let $\mathscr{P}_k = \{P(\cdot, n, \boldsymbol{\theta}), n \in N, \boldsymbol{\theta} \in \mathscr{H}_k\}$. Similarly, for $n \in N$ consider $a^n_{i_1 \cdots i_{k-1}} \geqslant 0$ and define analogously a class of probability measures $\mathscr{P}_{k-1}$ and continue this process to finally introduce $a^n_{i_1} \geqslant 0$ and define $\mathscr{P}_1$. The k-tuple $(\mathscr{P}_k, \mathscr{P}_{k-1},..., \mathscr{P}_1)$ is called a k-dimensional family. Thus if we take N as the set of positive integers and for each positive integer $n \in N$, let $a^n_{i_1 \cdots i_k} = n!/i_1! \cdots i_k! (n - i_1 - \cdots - i_k)!$ and define $\theta_i = p_i/\{1 - \sum_1^k p_i\}$, with $0 < p_i$, $i = 1,..., k$; $\sum_1^k p_i < 1$, then we get the k-dimensional family consisting of k-variate, $(k - 1)$-variate,..., trivariate, bivariate, and univariate positive multinomials.

Similarly, if, as before, we take N as the set of positive real numbers and for each n, let

$$a^n_{i_1 \cdots i_k} = \frac{(n + \sum_1^k i_j - 1)!}{i_1! \cdots i_k! (n - 1)!}$$

and define $\theta_i = a_i/(1 + \sum_1^k a_i)$, for $a_i > 0$, $i = 1,..., k$, we get the k-dimensional family consisting of k-variate, $(k - 1)$-variate,..., trivariate, bivariate, and univariate negative multinomials.

Hereafter the above two families of positive and negative multinomials as well as those obtained by linear transformations $Y_i = aX_i + b_i$, $i = 1,..., k$, are referred to as the family of multinomials. We can now state the properties enjoyed by the k-dimensional family of multinomials.

Let a k-dimensional family be given and let the joint distribution of $X_1,..., X_k$ belong to $\mathscr{P}_k$. Properties P1 to P4 may be described in the following way.

P1. The marginal distribution of $X_{i_1},..., X_{i_m}$ $(1 \leqslant i_1 < \cdots < i_m \leqslant k;$ $1 \leqslant m \leqslant k - 1)$ belongs to the class $\mathscr{P}_m$.

P2. The conditional distribution of $X_{i_1},..., X_{i_m}$ given $X_{j_1} = x_{j_1},..., X_{j_s} = x_{j_s}$ belongs to the class $\mathscr{P}_m$ and depends on the x_{j_i}'s only through $\sum_{l=1}^s x_{j_l}$, where $(i_1,..., i_m, j_1,..., j_s)$ are a permutation of $(1,..., k)$.

P3. The regression of X_i on $X_{i_1},..., X_{i_m}$ $(m \geqslant 1)$ is linear, where $(i, i_1,..., i_m) \subsetneq (1, 2,..., k)$.

P4. Define

$$Y_1 = \sum_{l=1}^{m_1} X_{i_l}, \quad Y_2 = \sum_{l=m_1+1}^{m_1+m_2} X_{i_l},..., \quad Y_s = \sum_{l=m_1+\cdots+m_{s-1}+1}^{m_1+\cdots+m_s} X_{i_l},$$

where $\sum_1^s m_i \leqslant k$ and $(i_1,..., i_{m_1+\cdots+m_s}) \subseteq (1,..., k)$. Then the joint distribution of $Y_1,..., Y_s$ is again a member of the class $\mathscr{P}_s$ and satisfies P1 to P3 with $k = s$.

In an earlier paper, Sinha and Sinha [6] utilized the properties relating to regression to characterize forms of nonsingular dispersion matrices of the underlying parent distributions. Interestingly enough, the only two types of

dispersion matrices so derived were found to correspond formally to those of multinomials. This opens up the possibility, in view of [3], that at least within the class of power series distributions these may be the only distributions characterized by these regression properties. It turns out, however, that is not true.

In fact we construct new bivariate (Sect. 3) and trivariate (Sect. 4) families satisfying all of Neyman's properties except part of P4 (namely, that the distribution of $X_1 + \cdots + X_k$ belongs to $\mathscr{P}_1$). One of the bivariate examples is used in the construction of the trivariate example. It may be noted that if all the properties listed by Neyman are to be meaningful, the dimension k must be greater than or equal to 3. These examples notwithstanding, it seems likely that within the class of power series distributions, the multinomials are characterized by the following properties.

Q1. The regression of X_i on the remaining variables is a linear function of the sum of the remaining variables.

Q2. The distribution of $X_1 + \cdots + X_k$ is of the power series type.

If true, this conjecture would certainly solve Neyman's problem since Q1 and Q2 obviously require less of $X_1, \ldots, X_k$ than P1 to P4. Under the condition that $P_\theta(X_1 + \cdots + X_k = 0)$ and $P_\theta(X_1 + \cdots + X_k = 1)$ are strictly positive the conjecture is proved in Section 5.

2. PRELIMINARIES

In this section we establish some properties of power series distribution that follow from linearity of regression. For simplicity we consider the case of three variables. Let X_1, X_2, and X_3 be three discrete random variables defined over I^3, having the power series $p \cdot f$.

$$P_\theta(X_1 = i, X_2 = j, X_3 = k) = a_{ijk}\theta_1{}^i\theta_2{}^j\theta_3{}^k/f(\theta), \qquad (2.1)$$

where $f(\theta) = \sum_{i,j,k} a_{ijk}\theta_1{}^i\theta_2{}^j\theta_3{}^k$, $a_{ijk} \geqslant 0$, $(i, j, k) \in I^3$, $\theta_1, \theta_2, \theta_3 > 0$. Let $\Omega = \{\theta \in R^3 : f(\theta) < \infty\}$. We assume in this section that $\mathscr{H}_3 \subset$ interior of Ω; this justifies the interchanges of differentiation and summation made in the proof of the proposition stated below. More generally, we could have assumed $\mathscr{H}_3 \subset$ the closure of the intersection of $\mathscr{H}_3$ and the interior of Ω; one could then first prove the proposition for $\mathscr{H}_3 \cap$ (interior of Ω) and then extend it to $\mathscr{H}_3$ by taking limits.

PROPOSITION. *Suppose that the following holds.*

$$E_{\theta_1}(X_1 \mid X_2 = j, \quad X_3 = k) = \alpha_1(\theta_1) + \beta_1(\theta_1)(j + k). \qquad (2.2)$$

Then (a) *the marginal bivariate distribution of* X_2 *and* X_3 *is of the power series type, and* (b) *the conditional distribution of* X_1, *given* $X_2 = j$, $X_3 = k$, *which is* (*trivially*) *of power series type, depends only on* $(j + k)$.

Proof. To prove (a), note that, by definition, (2.2) implies

$$\sum_i i a_{ijk}\theta_1{}^i \Big/ \sum_i a_{ijk}\theta_1{}^i = \alpha_1(\theta_1) + \beta_1(\theta_1)(j + k)$$

$$\Rightarrow \theta_1(d/d\theta_1) \log \left\{ \sum_i a_{ijk}\theta_1{}^i \right\} = \alpha_1(\theta_1) + \beta_1(\theta_1)(j + k)$$

$$\Rightarrow \sum_i a_{ijk}\theta_1{}^i = A(\theta_1) \cdot \{B(\theta_1)\}^{j+k} \cdot c(j, k) \qquad \text{for some} \quad A(\theta_1) > 0,$$
$$B(\theta_1) > 0, \quad c(j, k) \geqslant 0$$

$$\Rightarrow P_\theta(X_2 = j, X_3 = k) = \left(\sum_i a_{ijk}\theta_1{}^i \right) \theta_2{}^j \theta_3{}^k / f(\boldsymbol{\theta})$$

$$= c(j, k)\, \theta_2^{*j}\theta_3^{*k} / f^*(\boldsymbol{\theta}), \tag{2.3}$$

where $\theta_2{}^* = \theta_2 B(\theta_1)$, $\theta_3{}^* = \theta_3 B(\theta_1)$, $f^*(\boldsymbol{\theta}) = f(\boldsymbol{\theta})/A(\theta_1)$. This proves (a).

To prove (b), note that the generating function of the conditional univariate power series distribution of X_1 given $X_2 = j$, $X_3 = k$, defined by

$$\psi(\theta_1, s, j, k) = E_{\theta_1}(s^{X_1} \mid X_2 = j, X_3 = k)$$

satisfies

$$s \cdot (d/ds) \log \psi(\theta_1, s, j, k) = E_{\theta_1 s}(X_1 \mid X_2 = j, X_3 = k)$$
$$= \alpha_1(\theta_1 s) + \beta_1(\theta_1 s)(j + k) \tag{2.4}$$

by (2.2). Since $\psi(\theta_1, s, j, k) = 1$ for $s = 1$, (2.4) implies

$$\psi(\theta_1, s, j, k) = \phi(\theta_1, s, j + k),$$

thereby proving (b).

Remark. Note that a sort of partial converse of the proposition is valid in the sense that the validity of (2.4) implies that of (2.2). This can be proved using [2, p. 11, Lemma 1.1.2].

Note, further, that if X_1, X_2, and X_3 have a power series distribution as given in (2.1) then as noted by Patil [5, p. 184] and Bildikar and Patil [1, Theorem 4.1], the marginal distribution of X_2 and X_3 is of the form

$$P_\theta(X_2 = j, X_3 = k) = \theta_2{}^j \theta_3{}^k h(\theta_1, j, k)/f(\boldsymbol{\theta}),$$

which is of the power series type in our sense iff θ_1 is kept fixed or $h(\theta_1, j, k)$

can be expressed as $A(\theta_1)\{B(\theta_1)\}^j \{c(\theta_1)\}^k c(j, k)$. That the latter property is valid under linear regression follows from our proposition.

Suppose further that the marginal bivariate power series distribution (2.3) of X_2 and X_3 satisfies

$$E_{\theta_{2^*}}(X_2 \mid X_3 = k) = \alpha_2(\theta_2{}^*) + \beta_2(\theta_2{}^*)k. \tag{2.5}$$

An analysis similar to the one given above under (a) then immediately says that the marginal distribution of X_3 is again of the power series type.

3. Case of Two Variables

We first construct in this section a series of examples of joint discrete power series distributions of two variables X_1 and X_2, each with the values 0, 1, 2, and each having linear regression on the other. We take up a power series distribution

$$P_\theta(X_1 = i, X_2 = j) \propto a_{ij}\theta_1{}^i\theta_2{}^j; \qquad i, j = 0, 1, 2, \tag{3.1}$$

and assume that the distribution is symmetric, i.e., $a_{ij} = a_{ji} \geqslant 0$, $i, j = 0, 1, 2$, $\theta_1 > 0$, $\theta_2 > 0$. In order to exclude trivial cases, we subject them to

$$\sum_i a_{ij} > 0 \qquad \text{for each} \quad j = 0, 1, 2. \tag{3.2}$$

The regressions of X_1 on X_2 are given by the expressions

$$E(X_1 \mid X_2 = j) = \{a_{j1}\theta_1 + 2a_{j2}\theta_1{}^2\}/\{a_{j0} + a_{j1}\theta_1 + a_{j2}\theta_1{}^2\}, \qquad j = 0, 1, 2. \tag{3.3}$$

If the regression is linear, i.e., $E_{\theta_1}(X_1 \mid X_2 = j) = \alpha(\theta_1) + \beta(\theta_1)j$, then (3.3) leads to three equations in α and β. Eliminating α and β, we find that the consistency requirement connecting the a_{ij}'s is given by

$$E(X_1 \mid X_2 = 2) = 2E(X_1 \mid X_2 = 1) - E(X_1 \mid X_2 = 0) \tag{3.4}$$

identically in θ_1, which yields the equations

$$2a_{00}a_{02}a_{11} = a_{01}(a_{00}a_{12} + a_{01}a_{02}), \tag{3.4.1}$$

$$2a_{01}(a_{00}a_{22} + a_{01}a_{12} + a_{02}^2) = a_{11}(a_{00}a_{12} + a_{01}a_{02}) + 4a_{00}a_{02}a_{12}, \tag{3.4.2}$$

$$a_{00}a_{12}^2 = a_{01}^2a_{22}, \tag{3.4.3}$$

$$4a_{01}a_{02}a_{22} + a_{11}(a_{12}a_{02} + a_{01}a_{22}) = 2a_{12}(a_{00}a_{22} + a_{02}^2 + a_{01}a_{12}), \tag{3.4.4}$$

$$2a_{02}a_{11}a_{22} = a_{12}(a_{01}a_{22} + a_{12}a_{02}). \tag{3.4.5}$$

We first indicate the nature of the solutions involving some of the a_{ij}'s as zeros.

(i) $a_{00} = a_{01} = 0$, $a_{11} = 2a_{02}$, $2a_{11}a_{22} = a_{12}^2$, $a_{02} > 0$, $a_{12} > 0$.

(ii) $a_{12} = a_{22} = 0$, $a_{02} = a_{01}^2/4a_{00}$, $a_{11} = a_{01}^2/2a_{00}$, $a_{00} > 0$, $a_{01} > 0$ (yielding the trinomial distribution of X_1, X_2, and $2 - X_1 - X_2$).

(iii) $a_{01} = a_{02} = a_{12} = 0$, $a_{00} > 0$, $a_{11} > 0$, $a_{22} > 0$ (yielding yet another trivial solution).

Next assume all a_{ij}'s are positive. We may add one conventional equation

$$\sum_i \sum_j a_{ij} - 1. \tag{3.5}$$

It is easy to observe now that (3.4.3) together with (3.4.1) implies (3.4.5), and (3.4.3) together with (3.4.2) implies (3.4.4). Hence, only (3.4.1) to (3.4.3) and (3.5) are to be considered. Now

(3.4.3) $\Rightarrow a_{12} = a_{01}(a_{22}/a_{00})^{1/2}$,

(3.4.1) $\Rightarrow a_{11} = \{a_{01}^2/2a_{00}a_{02}\}\{a_{02} + (a_{00}a_{22})^{1/2}\}$,

(3.4.2) $\Rightarrow 2a_{01}\{a_{00}a_{22} + a_{01}^2(a_{22}/a_{00})^{1/2} + a_{02}^2\}$

$\qquad = \{a_{01}^3/2a_{00}a_{02}\}(a_{02} + (a_{00}a_{22})^{1/2})^2 + 4a_{00}a_{02}a_{01}(a_{22}/a_{00})^{1/2};$

i.e.,

$$2(a_{02} - (a_{00}a_{22})^{1/2})^2 + 2a_{01}^2(a_{22}/a_{00})^{1/2} = \{a_{01}^2/2a_{00}a_{02}\}\{a_{02} + (a_{00}a_{22})^{1/2}\}^2.$$

This yields a cubic equation in a_{02} of which the solutions are $a_{02} = a_{01}^2/4a_{00}$ and $a_{02} = (a_{00}a_{22})^{1/2}$ (with multiplicity 2). For $a_{02} = a_{01}^2/4a_{00}$, we have $a_{11} = 2\{(a_{01}^2/4a_{00}) + (a_{00}a_{22})^{1/2}\}$, $a_{12} = (a_{01}/a_{00})(a_{00}a_{22})^{1/2}$ while from (3.5), we have, further,

$$a_{22} = \{1 - (a_{00})^{1/2} - a_{01}/(a_{00})^{1/2}\}^2.$$

Hence, one solution is

$$\left.\begin{array}{l} a_{00} \, (>0) \\ a_{01} \, (>0) \end{array}\right\} \quad \text{subject to} \quad (a_{00})^{1/2} + a_{01}/(a_{00})^{1/2} < 1;$$

$$a_{02} = a_{01}^2/4a_{00},$$

$$a_{22} = (1 - (a_{00})^{1/2} - a_{01}/(a_{00})^{1/2})^2,$$

$$a_{12} = a_{01}/a_{00} \cdot (a_{00}a_{22})^{1/2},$$

$$a_{11} = a_{01}^2/2a_{00} + 2(a_{00}a_{22})^{1/2}. \tag{3.6}$$

Set $a_{00} = p_{00}^2$ and $a_{01} = 2p_{01}p_{00}$ with $p_{00} + 2p_{01} < 1$. Then the solution becomes

$$a_{00} = p_{00}^2, \qquad a_{01} = 2p_{00}p_{01}, \qquad a_{02} = p_{01}^2, \qquad a_{22} = p_{11}^2, \tag{3.6'}$$
$$a_{12} = 2p_{01}p_{11}, \qquad a_{11} = 2(p_{01}^2 + p_{00}p_{11}),$$

where $p_{00} + 2p_{01} + p_{11} = 1$.

This solution can be identified as the one generated as a twofold convolution of the joint discrete distribution whose generating function is

$$G(t_1, t_2) = (p_{00} + p_{01}t_2 + p_{10}t_1 + p_{11}t_1t_2) \qquad \text{with} \quad p_{01} = p_{10}.$$

In this connection see Remark 1 at the end of Section 4.

Another solution is obtained by setting $a_{02} = (a_{00}a_{22})^{1/2}$ and here we get

$$\left. \begin{array}{l} a_{00}(>0) \\ a_{01}(>0) \end{array} \right\} \quad \text{subject to} \quad (a_{00})^{1/2} + a_{01}/(a_{00})^{1/2} < 1;$$
$$a_{02} = (a_{00}a_{22})^{1/2},$$
$$a_{11} = a_{01}^2/a_{00}, \tag{3.7}$$
$$a_{12} = a_{01}(a_{22}/a_{00})^{1/2},$$
$$a_{22} = (1 - (a_{00})^{1/2} - a_{01}/(a_{00})^{1/2})^2.$$

Thus we have succeeded in characterizing all possible forms of a bivariate discrete power series distribution of two variables X_1 and X_2 each with the values 0, 1, 2 and exhibiting linear regression of X_1 on X_2 and of X_2 on X_1 (due to symmetry).

Let $a_{ij}\theta_1^i\theta_2^j/c(\theta_1, \theta_2)$ be one such distribution. Observe that, in view of the results of Section 2, the linear regression of $X_2(X_1)$ on $X_1(X_2)$ implies that the marginal distribution of $X_1(X_2)$ is of the power series type. Let $N = \{1, 2, 3, 4, 5\}$, $a_{ij}^n = a_{ij}$, $n = 1$ to 5, and as in Section 1, define $\mathscr{P}_2$ as $\{P(\cdot, n, \theta), n \in N, \theta \in \mathscr{H}_2\}$. If we now consider the univariate marginal power series distributions of X_1 and X_2, and the conditional univariate power series distributions of X_1 given $X_2 = 0, 1,$ and 2, we get a collection of five univariate power series distributions. These five distributions indexed arbitrarily by $n \in N$ will constitute $\mathscr{P}_1$. Since $a_{ij} = a_{ji}$, $i, j = 0, 1, 2$, it is clear that $\{\mathscr{P}_1, \mathscr{P}_2\}$ satisfies the properties P1, P2, and P3.

4. Case of Three Variables

As in Section 3, we first construct an example of a joint discrete power series distribution of three variables X_1, X_2, and X_3, each with the values 0, 1, 2, such that

$$E_{\theta_1}(X_1 \mid X_2 = j, \quad X_3 = k) = \alpha_1(\theta_1) + \beta_1(\theta_1)(j + k),$$

$$E_{\theta_1,\theta_3}(X_1 \mid X_2 = j) = \alpha_2(\theta_1, \theta_3) + \beta_2(\theta_1, \theta_3)j,$$

$$E_{\theta_1,\theta_2}(X_1 + X_2 \mid X_3 = k) = \alpha_3(\theta_1, \theta_2) + \beta_3(\theta_1, \theta_2)k, \tag{4.1}$$

$$E_{\theta_1}(X_1 \mid X_2 + X_3 = l) = \alpha_4(\theta_1) + \beta_4(\theta_1)l.$$

Let

$$P_\theta(X_1 = i, X_2 = j, X_3 = k) = P_{ijk} \propto a_{ijk}\theta_1{}^i\theta_2{}^j\theta_3{}^k, \tag{4.2}$$

where $a_{ijk} \geqslant 0$, a_{ijk} is symmetric in the arguments i, j, k, $\theta_1 > 0$, $\theta_2 > 0$, $\theta_3 > 0$, and moreover, to exclude trivial cases,

$$\sum_j \sum_k a_{ijk} > 0 \qquad \text{for each} \quad i = 0, 1, 2. \tag{4.3}$$

We utilize some of the results of Section 3 to derive proper relations connecting the a_{ijk}'s so as to satisfy (4.1). Toward this end, we first write the marginal bivariate distribution of (X_1, X_2) as

$$P_\theta[X_1 = i, X_2 = j] = P_{ij}^* \propto A_{ij}\theta_1{}^i\theta_2{}^j, \tag{4.4}$$

where $A_{ij} = (a_{ij0} + \theta_3 a_{ij1} + \theta_3{}^2 a_{ij2})$; $i, j = 0, 1, 2$. Now note that (4.4) is required to satisfy the property of linearity of regression of X_1 on X_2. This, however, immediately reveals all possible relations connecting the A_{ij}'s as discussed in Section 3. In particular, we take up solution (i) with some of the A_{ij}'s as zeros. Thus, we get

(a) $A_{00} = A_{01} = 0$ identically in $\theta_3 > 0$, which gives

$$a_{000} = a_{001} = a_{002} = 0,$$

$$a_{011} = a_{012} = 0. \tag{4.5}$$

(b) $A_{11} = 2A_{02}$. However, in view of (4.5), this yields $a_{111}\theta_3 + a_{112}\theta_3{}^2 = 2a_{022}\theta_3{}^2$ identically in $\theta_3 > 0$ and hence we derive

$$a_{111} = 0,$$

$$a_{112} = 2a_{022}. \tag{4.6}$$

(c) $2A_{11}A_{22} = A_{12}^2$. This again, in view of (4.5) and (4.6), yields

$$2(a_{112}\theta_3{}^2)(a_{022} + a_{122}\theta_3 + a_{222}\theta_3{}^2) = (a_{112}\theta_3 + a_{122}\theta_3{}^2)^2$$

identically in $\theta_3 > 0$, from which we derive only one more equation,

$$2a_{112}a_{222} = a_{122}^2.$$

Thus, finally, we may conclude that a choice of the a_{ijk}'s such as

$$a_{000} = a_{001} = a_{002} = a_{011} = a_{012} = a_{111} = 0, \qquad a_{022} > 0, \qquad a_{112} > 0, \tag{4.7}$$

$$a_{122} > 0, \qquad a_{222} > 0, \qquad a_{112} = 2a_{022}, \qquad 2a_{112}a_{222} = a_{122}^2$$

will lead to linear regression of any variable on any other variable in their bivariate marginal distribution (because of symmetry of the a_{ijk}'s). It is easy to verify that this solution is nontrivial in the sense that there do not exist multinomial variables Y_1, Y_2, and Y_3, and two constants a and b such that $X_i = aY_i + b, i = 1, 2, 3$. Now we work out the regression of X_1 on X_2 and X_3. For this, we prepare the following table of $P\{X_1 = i \mid X_2 = j, X_3 = k\}$, using (4.7).

j	k	i	Conditional probability of X_1 : $P\{X_1 = i \mid X_2 = j, X_3 = k\}$
0	2	2	1
1	1	2	1
1	2	$(1, 2)$	$\left(\dfrac{P_{112}}{P_{112} + P_{212}}, \dfrac{P_{212}}{P_{112} + P_{212}}\right)$
2	2	$(0, 1, 2)$	$\left(\dfrac{P_{022}}{P_{022} + P_{122} + P_{222}}, \dfrac{P_{122}}{P_{022} + P_{122} + P_{222}}, \dfrac{P_{222}}{P_{022} + P_{122} + P_{222}}\right)$

It is now easy to verify that with

$$\alpha_1(\theta_1) = \frac{2\{2a_{112} + \theta_1 a_{122}\}}{a_{112} + \theta_1 a_{122}}$$

and $\beta_1(\theta_1) = -a_{12}/\{a_{112} + \theta_1 a_{122}\}$,

$$E_{\theta_1}(X_1 \mid X_2 = j, X_3 = k) = \alpha_1(\theta_1) + \beta_1(\theta_1)(j + k),$$

in view of $a_{112} = 2a_{022}$ and $2a_{112}a_{222} = a_{122}^2$, as listed in (4.7). Thus, a choice of the a_{ijk}'s such as (4.7) will meet the first two conditions under (4.1) and hence all the conditions under (4.1). Consequently, in view of the results of Section 2, the marginal bivariate distribution of any two variables is of power series type, the marginal univariate distribution of any single variable is of power series type, and, trivially, the conditional bivariate distribution of any two variables given the third is of power series type and the conditional univariate distribution of any single variable given the other two is also of power series type. Also, direct computation shows that the bivariate distribution of $Y_1 = X_1 + X_2$ and $Y_2 = X_3$ is given by

$$P_\theta(Y_1 = i, Y_2 = j) \propto A_{ij}^* \theta^{*i} \theta_3^{\,j},$$

where $\theta^* = \theta_1\theta_2/(\theta_1 + \theta_2)$, $A_{22}^* = a_{022}$, $A_{31}^* = a_{112}$, $A_{32}^* = a_{122}$, $A_{40}^* = a_{022}$,

$A_{41}^* = a_{122}$, and $A_{42}^* = a_{222}$. This is certainly of the same form as (4.2), except for symmetry of the A_{ij}^*'s, which we do not need here.

If we now start with any power series distribution satisfying (4.1) and construct $(\mathscr{P}_1, \mathscr{P}_2, \mathscr{P}_3)$ in the same way as in the last paragraph of Section 3, then $(\mathscr{P}_1, \mathscr{P}_2, \mathscr{P}_3)$ will have all the properties P1 to P4 except for the part of P4 relating to the distribution of $X_1 + X_2 + X_3$.

Remark 1.   Let $\{X_1^{(i)}, X_2^{(i)}, X_3^{(i)}\}$, $i = 1,..., N$, be independent copies of a triplet (X_1, X_2, X_3) satisfying (4.1). It is then easy to verify that the distribution of the N-fold convolution (Z_1, Z_2, Z_3) where $Z_j = \sum_{i=1}^{N} X_j^{(i)}$, $j = 1, 2, 3$, also satisfies (4.1). We can then use this to construct new families satisfying Neyman's properties except for the part of P4 relating to the distribution of the sum $Z_1 + Z_2 + Z_3$.

Remark 2.   Let Y_1, Y_2, and Y_3 be independent Poisson variables with the means $\lambda_1\theta_1$, $\lambda_2\theta_2$, and $\lambda_3\theta_1\theta_2$, respectively ($\lambda$'s fixed). Define now the two random variables $X_1 = Y_1 + Y_3$, $X_2 = Y_2 + Y_3$. It is easy to verify that the bivariate discrete distribution of X_1 and X_2 satisfies the property of having linear regression of one variable on the other. We can therefore construct a bivariate family satisfying Neyman's properties except, again, for the part of P4 relating to the distribution of the sum of the variables.

5. CHARACTERIZATION OF THE MULTINOMIALS

Let $X_1,..., X_k$ have a joint k-variate discrete power series distribution with the p.f.

$$P_\theta(X_1 = i_1,..., X_k = i_k) = a_{i_1 \cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k}/f(\theta) \qquad (5.1)$$

where $a_{i_1 \cdots i_k} \geq 0$, $(i_1,..., i_k) \in I^k$, $\theta_1,..., \theta_k > 0$, and

$$f(\theta) = \sum_{i_1,...,i_k} a_{i_1 \cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k}.$$

We assume $|f(\theta)| < \infty$ $\forall\theta$ in a sufficiently small neighborhood of the origin. This justifies (5.5), (5.6), and (5.7), at least for sufficiently small θ, which is all that is needed in the proof of Theorem 5.1 stated below.

THEOREM 5.1.   *Under the conditions $a_{0\cdots 0} > 0$ and $a_{10\cdots 0} + \cdots + a_{0\cdots 01} > 0$, properties* Q1 *and* Q2 *characterize the multinomials.*

Proof.   Obviously property Q2 implies

$$\sum_{i_1+\cdots+i_k=m} a_{i_1\cdots i_k}\theta_1^{i_1} \cdots \theta_k^{i_k}/f(\theta) = b(m)\{c(\theta)\}^m A(\theta), \qquad m \geq 0. \quad (5.2)$$

For $m = 0$ and 1, (5.2) implies, under the assumed conditions, $f(\theta) A(\theta) = a_{0\ldots0}/b(0)$ and $c(\theta) = \sum_1^k c_i \theta_i = \sum_1^k \theta_i'$(say) with $\theta_i' = c_i \theta_i$, $i = 1,\ldots, k$, and $c_1 = a_{10\ldots0} b(0)/a_{0\ldots0} b(1),\ldots,$ $c_k = a_{0\ldots01} b(0)/a_{0\ldots0} b(1)$. Reparametrizing and denoting θ_i''s by θ_i's, we can write (5.1) as

$$P_\theta(X_1 = i_1,\ldots, X_k = i_k) = b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k}/f^*(\theta) \qquad (5.3)$$

so that Q2 will imply

$$\sum_{i_1+\cdots+i_k=m} b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k} = c^* b(m) \left(\sum_1^k \theta_i\right)^m \text{ for some constant } c^* > 0$$

$$\Rightarrow \sum_{i_1,\ldots,i_k} b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k} = c^* \sum_m b(m) \left(\sum_1^k \theta_i\right)^m$$

$$= c^* \psi \left(\sum_1^k \theta_i\right) \qquad \text{(say).} \qquad (5.4)$$

Note that (5.4) implies

$$\partial\psi/\partial\theta_1 = \partial\psi/\partial\theta_i, \qquad \text{for all } i \neq 1. \qquad (5.5)$$

Now property Q1 implies

$$\sum_{i_1} i_1 b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k} \Big/ \sum_{i_1} b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k} = \alpha(\theta_1) + \beta(\theta_1)(i_2 + \cdots + i_k)$$

$$\Rightarrow \sum_{i_1,\ldots,i_k} i_1 b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k} = \alpha(\theta_1) \sum_{i_1,\ldots,i_k} b_{i_1\cdots i_k} \theta_1^{i_1} \cdots \theta_k^{i_k}$$

$$+ \beta(\theta_1) \sum_{i_1,\ldots,i_k} (i_2 + \cdots + i_k) b_{i_1\cdots i_k}$$

$$\times \theta_1^{i_1} \cdots \theta_k^{i_k}$$

$$\Rightarrow \left[\theta_1 - \beta(\theta_1) \sum_2^k \theta_i\right] \cdot (1/\psi)(\partial\psi/\partial\theta_1) = \alpha(\theta_1), \qquad \text{using (5.4) and (5.5)}$$

$$\Rightarrow \eta\left(\sum_1^k \theta_i\right) = \alpha(\theta_1) \Big/ \left(\theta_1 - \beta(\theta_1) \sum_2^k \theta_i\right), \qquad (5.6)$$

where $\eta = (1/\psi) \cdot (\partial\psi/\partial\theta_1)$. We now use the fact that $\partial\eta/\partial\theta_1 = \partial\eta/\partial\theta_i$ for any $i \neq 1$. This gives, in view of (5.6),

$$\alpha'(\theta_1) \left[\theta_1 - \beta(\theta_1) \sum_2^k \theta_i\right] - \alpha(\theta_1) \left[1 - \beta'(\theta_1) \sum_2^k \theta_i\right] = \alpha(\theta_1)\beta(\theta_1)$$

$$\Rightarrow \alpha'(\theta_1)\theta_1 - \alpha(\theta_1) = \alpha(\theta_1)\beta(\theta_1) \qquad \text{and} \qquad \alpha(\theta_1)\beta'(\theta_1) - \alpha'(\theta_1)\beta(\theta_1) = 0$$

$$\Rightarrow d/d\theta_1 \left[\alpha(\theta_1)/\beta(\theta_1)\right] = 0 \quad \text{and} \quad (d/d\theta_1)[\theta_1/\alpha(\theta_1)] = -\beta(\theta_1)/\alpha(\theta_1)$$

$$\Rightarrow \beta(\theta_1) = c\alpha(\theta_1) \quad \text{and} \quad \theta_1/\alpha(\theta_1) = -c\theta_1 + d$$

for some constants c and d. Therefore, (5.6) reduces to

$$\eta \left(\sum_1^k \theta_i \right) = 1 \Big/ \left(d - c \sum_1^k \theta_i \right) ;$$

i.e.,

$$(d/d\xi) \log \psi(\xi) = 1/(d - c\xi), \qquad \xi = \sum_1^k \theta_i . \tag{5.7}$$

Since $a_{0\cdots 0} > 0$, $\log \psi$ is analytic at the origin. This and (5.7) imply $d \neq 0$. So

$$\psi(\xi) = L \cdot (1 + (\xi/\hat{c}d))^{\hat{c}} \quad \text{for} \quad d \neq 0, \tag{5.8}$$

where $\hat{c} = -1/c$ and L is a constant (independent of ξ). Since ψ is the generating function of $b_{i_1 \cdots i_k}$, which are nonnegative, $L > 0$, $d > 0$, and $\hat{c}$ is either a positive integer or a negative real number. From (5.8) one can get the $b_{i_1 \cdots i_k}$'s and hence the $a_{i_1 \cdots i_k}$'s in terms of L, d, and $\hat{c}$. It is trivial to check that (5.1) reduces to a positive or negative multinomial according as $\hat{c}$ is a positive integer or a negative real number. This completes the proof.

Remark. Our examples in Sections 3 and 4 show that Q1 individually does not lead to a characterization of the multinomials. It is easy to construct examples to show that Q2 also individually does not do so.

The main problem now is to characterize all power series distributions which satisfy Q1 and Q2. As a generalization of this we may, as in [1], replace power series distributions by exponential distributions in Q2 and try to characterize all exponential distributions having properties Q1 and Q2. It is clear that the multivariate normal, with the mean vector taking arbitrary values and a fixed dispersion matrix satisfying the conditions of Sinha and Sinha [6] is an example of this sort. Are the multinomials and the multivariate normal of the above type the only examples?

REFERENCES

[1] BILDIKAR, S. AND PATIL, G. P. (1968). Multivariate exponential-type distributions. *Ann. Math. Statist.* **39** 1316–1326.
[2] KAGAN, A. M., LINNIK, YU. V. AND RAO, C. R. (1973). *Characterization Problems in Mathematical Statistics.* Wiley, New York.
[3] KHATRI, C. G. (1959). On certain properties of power-series distributions. *Biometrika* **46** 486–490.
[4] NEYMAN, J. (1965). Certain chance mechanisms involving discrete distributions. In *Classical and Contagious Discrete Distributions* (G. P. Patil, Ed.), pp. 4–14.
[5] PATIL, G. P. (1965). On multivariate generalized power series distribution and its application to the multinomial and negative multinomial. In *Classical and Contagious Discrete Distributions* (G. P. Patil, Ed.), pp. 183–194.
[6] SINHA, B. K., AND SINHA, B. K. (1976). On a characterization of the dispersion matrix based on the properties of regression. *Comm. Statist. A* **5** (13) 1215–1224.