On some bivariate discrete distributions with multivariate components

By JOHN PANARETOS (Columbia, Missouri) *)

Abstract

Consider a random vector $(\underset{\sim}{X}, \underset{\sim}{Y})$ where $\underset{\sim}{X}=(X_1, X_2, \ldots, X_s)$ and $\underset{\sim}{Y}=(Y_1, Y_2, \ldots, Y_s)$ with X_i, Y_i, $i=1, 2, \ldots, s$ independent non-negative, integer-valued random variables with finite support and such that $\underset{\sim}{X} \geqq \underset{\sim}{Y}$. We show that in the case where the distribution of $(\underset{\sim}{Y}|\underset{\sim}{X}=\underset{\sim}{n})$ is of a certain structural form then there exists a relationship between the distributions of $\underset{\sim}{Y}$ and of $\underset{\sim}{Y}|(\underset{\sim}{X}=\underset{\sim}{Y})$ which uniquely determines the distribution of $\underset{\sim}{X}$. The relationship in question is less stringent that the condition of independence between $\underset{\sim}{Y}$ and $\underset{\sim}{X}-\underset{\sim}{Y}$ usually involved in problems of this nature. Examples are given to illustrate the result. The case where $\underset{\sim}{X}, \underset{\sim}{Y}$ heve infinite support has been examined earlier by the author.

Keywords and Phrases. Finite Distributions, Conditional Distribution, Multiple Binomial Distribution, Multiple Hypergeometric Distribution, Characterization.

AMS 1980 Subject Classification: Primary: 62H05, Secondary: 60E05.

1. Introduction

Let (X, Y) be a random vector with non-negative, integer-valued components with finite support and assume that $X \geqq Y$. In a recent paper [4] the author studied the problem of finding the distribution of X (and hence the joint distribution of (X, Y)) when the conditional distribution of Y on X was of the structural from

$$P(Y=r|X=n) = \frac{a_r b_{n-r}}{c_n} \quad r = 0, 1, \ldots, n; \quad n = 0, 1, \ldots, N; \quad N > 0.$$

It was pointed out that independence between Y and $X-Y$ was not necessary in order that the form of the distribution of X be uniquely determined. In fact it was shown that a relationship between Y and $X-Y$ less stringent than independence was sufficient to determine the distribution of X. (This has now been extended (PANARETOS [6] to a very general case where no assumption is made regarding the structural form of the distribution of $Y|(X=n)$). The paper extended similar results by SHANBHAG [9] and PANARETOS [3] dealing with the same problem, but for distributions with infinite support.

*) Now at the University of Crete, Greece

In this paper we provide a multivariate analogue of Panaretos's [4] result in the sense that we take $\underset{\sim}{X}, \underset{\sim}{Y}$ to be random vectors with independent components. One may argue that in practice the assumption of independent marginal distributions is unlikely to be satisfied exactly. There are, however, situations where for reasons of mathematical simplicity one considers such distributions as a first approximation. This approach has the advantage of making the problem mathematically tractable.

In the next section of the paper we re-state the result in the univariate case for ease of reference. We then state and prove the multivariate extension of it. Then in Section 3 two examples are given as an illustration where the multiple binomial and the multiple hypergeometric distributions are characterized.

2. The Main result

Theorem 1. (PANARETOS, [4].) *Let* $\{(a_n, b_n)\colon n=0, 1, \ldots\}$ *be a sequence of nonnegative real vectors such that* $a_n>0$, $n=0, 1, \ldots, m$; $(m>0)$; $b_0>0$; $b_{jm+1}>0$, $j=0, 1, \ldots, k-1$; $0<k\leqq\left[\frac{N-1}{m}\right]$; $N>m$. *([Z] denotes the integral part of Z). Let* $\{c_n\colon n=0, 1, \ldots, N\}$ *denote the sequence* $\left\{\sum_{r=0}^{n} a_r b_{n-r}\colon n=0, 1, \ldots, N\right\}$. *Consider a random vector* (X, Y) *of nonnegative, integer-valued components such that* $P(X=n)=g_n$, $n=0, 1, \ldots, N$ *with* $g_0<1$ *and* $X\geqq Y$. *Let* $Z=X-Y$. *Suppose that whenever* $g_n>0$

$$(2.1) \qquad P(Y=r|X=n) = \frac{a_r b_{n-r}}{c_n}, \quad r=0, 1, \ldots, n; \quad n=0, 1, \ldots, N$$

Then

$$(2.2) \qquad P(Y=r|Z=0) = P(Y=r|Z=1) = P(Y=r|Z=m+1) =$$

$$= \ldots = P(Y=r|Z=(k-1)m+1) 1 \leqq k \leqq \left[\frac{N-1}{m}\right]$$

(k relations) if and only if (iff)

$$g_n = g_0 \frac{c_n}{c_0} \theta^n \quad \textit{for some} \quad \theta>0$$

$$(2.3) \qquad n=0, 1, \ldots, km+1, \quad 1 \leqq k \leqq \left[\frac{N-1}{m}\right]$$

The multivariate analogue of Theorem 1 can be stated in the following form.

Theorem 2. *Consider a random vector* $(\underset{\sim}{X}, \underset{\sim}{Y})$ *with* $\underset{\sim}{X}=(X_1, X_2, \ldots, X_S)$ *and* $\underset{\sim}{Y}=(Y_1, Y_2, \ldots, Y_s)$ *where* X_i, Y_i; $i=1, 2, \ldots, s$ *are non-negative integer-valued random variables such that* $P(\underset{\sim}{X}=\underset{\sim}{n})\equiv P(X_1=n_1, X_2=n_2, \ldots, X_s=n_s)=g_{\underset{\sim}{n}}$; $\underset{\sim}{n}=(n_1, n_2, \ldots, n_s)$; $n_i=0, 1, \ldots, N_i$; $N_i>0$; $i=1, 2, \ldots, s$ *with* $g_{\underset{\sim}{0}}<1$ *and* $\underset{\sim}{X}\geqq\underset{\sim}{Y}$. *Denote by* $\underset{\sim}{Z}=\underset{\sim}{X}-\underset{\sim}{Y}$. *Let* $\{(a_{\underset{\sim}{n}}, b_{\underset{\sim}{n}})\colon \underset{\sim}{n}=(n_1, n_2, \ldots, n_s)$; $n_i=0, 1, \ldots$; $i=1, 2, \ldots, s\}$ *be a sequence of real vectors such that* $a_{\underset{\sim}{r}}>0$, $r_i=0, 1, \ldots, m_i$, *for some* m_i *in*

$(0, N_i)$, $i=1,2,\ldots,s$; $b_{\underset{\sim}{n}}>0$ *for all* $\underset{\sim}{n}: n_i=0,1,m_i+1, 2m_i+1, \ldots, (k_i-1)m_i+1$; $0<k_i\leqq\left[\frac{N_i-1}{m_i}\right]$; $i=1,2,\ldots,s$. *Define the sequence* $\{c_{\underset{\sim}{n}}\}$ *to be:* $c_{\underset{\sim}{n}}=\sum_{\underset{\sim}{r}=\underset{\sim}{0}}^{\underset{\sim}{n}} a_{\underset{\sim}{r}} b_{\underset{\sim}{n}-\underset{\sim}{r}}$ $\underset{\sim}{n}=(n_1,n_2,\ldots,n_s)$; $n_i=0,1,\ldots,N_i$; $i=1,2,\ldots,s$ *where* $a_{\underset{\sim}{r}}\equiv(a_{r_1,r_2,\ldots,r_s})$ *and* $\sum_{\underset{\sim}{r}=\underset{\sim}{0}}^{\underset{\sim}{n}}$ *stands for* $\sum_{r_1=0}^{n_1}\sum_{r_2=0}^{n_2}\cdots\sum_{r_s=0}^{n_s}$. *Suppose that whenever* $g_n>0$

$$P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{X}=\underset{\sim}{n})=\frac{a_{\underset{\sim}{r}}b_{\underset{\sim}{n}-\underset{\sim}{r}}}{c_{\underset{\sim}{n}}} \tag{1.4}$$

$$r_i=0,1,\ldots,n_i;\quad n_i=0,1,\ldots,N_i;\quad i=1,2,\ldots,s.$$

Let $\{h_1,h_2,\ldots,h_w\}$, $w\leqq s$ *denote any subset of size* w *of* $\{1,2,\ldots,s\}$. *Then the conditions*

$$\begin{aligned}(1.5)\qquad P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{Z}=\underset{\sim}{0}) = P(\underset{\sim}{Y}=\underset{\sim}{r}|Z_{h_1}=l_{h_1}m_{h_1}+1, Z_{h_2}=l_{h_2}m_{h_2}+1,\ldots,\\ Z_{h_j}=l_{h_j}m_{h_j}+1, Z_{h_{j+1}}=0, Z_{h_{j+2}}=0,\ldots,Z_{h_s}=0)\end{aligned}$$

for all $j=1,2,\ldots,s$ *and all* $l_i=0,1,\ldots,k_i-1$; $1\leqq k_i\leqq\left[\frac{N_i-1}{m_i}\right]$;

$$r_i=0,1,\ldots,m_i;\quad i=1,2,\ldots,s$$

hold iff

$$g_{\underset{\sim}{n}}=g_{\underset{\sim}{0}}\frac{c_{\underset{\sim}{n}}}{c_{\underset{\sim}{0}}}\prod_{i=1}^{s}\theta_i^{n_i} \tag{1.6}$$

for some $\theta_i>0$; $i=1,2,\ldots,s$ *and all* $\underset{\sim}{n}=(n_1,n_2,\ldots,n_s)$ *such that* $n_i\leqq k_im_i+1$; $1\leqq k_i\leqq\left[\frac{N_i-1}{m_i}\right]$; $i=1,2,\ldots,s$.

Proof. (We prove the theorem for $s=2$. The case for $s>2$ then follows easily.)

"*If*" *Part:* From (1.4), (1.6), for $s=2$ and for given j_1 and j_2 such that $0\leqq j_i\leqq N_i$, $i=1,2$ we have

$$P(\underset{\sim}{Y}=\underset{\sim}{r}|Z_1=j_1,\ Z_2=j_2)=\frac{P(\underset{\sim}{Y}=\underset{\sim}{r},\ X_1=r_1+j_1,\ X_2=r_2+j_2)}{P(\underset{\sim}{Z}=\underset{\sim}{j})}=$$

$$=\frac{\dfrac{a_{\underset{\sim}{r}}b_{\underset{\sim}{j}}}{c_{\underset{\sim}{r}+\underset{\sim}{j}}}g_{\underset{\sim}{r}+\underset{\sim}{j}}}{P(\underset{\sim}{Z}=\underset{\sim}{j})}=\frac{a_{\underset{\sim}{r}}b_{\underset{\sim}{j}}}{P(\underset{\sim}{Z}=\underset{\sim}{j})}\frac{g_{\underset{\sim}{0}}}{c_{\underset{\sim}{0}}}\theta_1^{r_1+j_1}\theta_2^{r_2+j_2}$$

$$r_i=0,1,\ldots;\quad i=1,2.$$

Hence

$$P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{Z}=\underset{\sim}{j})=\frac{a_{\underset{\sim}{r}}\theta_1^{r_1}\theta_2^{r_2}}{\varphi(\underset{\sim}{j},\underset{\sim}{\theta})}\quad r_i=0,1,\ldots;\quad i=1,2.$$

We have, however, that $\sum_{\underset{\sim}{r}} P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{Z}=\underset{\sim}{j})=1$. So,

$$\varphi(\underset{\sim}{j}, \underset{\sim}{\theta}) = \sum_{\underset{\sim}{r}} P(\underset{\sim}{Y} = \underset{\sim}{r}|\underset{\sim}{Z} = \underset{\sim}{j})\, \theta_1^{r_1} \theta_2^{r_2} = A(\underset{\sim}{\theta}).$$

Hence

$$P(\underset{\sim}{Y} = \underset{\sim}{r}|\underset{\sim}{Z} = \underset{\sim}{j}) = \frac{a_{\underset{\sim}{r}}\, \theta_1^{r_1} \theta_2^{r_2}}{A(\underset{\sim}{\theta})}. \tag{1.7}$$

Note that j_1, j_2 were fixed but arbitrary. So (1.7) is valid for any j_1, j_2 such that $0 \leqq j_i \leqq N_i$ $i=1, 2$. The fact that the right-hand-side of (1.7) does not depend on $\underset{\sim}{j}$ for $j_i=0, 1, \dots, N_i$, $i=1, 2$ implies that the probabilities $P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{Z}=\underset{\sim}{j})$ are independent of $\underset{\sim}{j}$ and therefore they are all equal for a given $\underset{\sim}{r}$. This establishes (1.5).

Note. One may observe that if (1.4) and (1.6) are valid then $\underset{\sim}{Y}$ and $\underset{\sim}{X}-\underset{\sim}{Y}$ not only satisfy (1.5), but they are in fact completely independent.

"*Only if*" *Part*.

From $P(\underset{\sim}{Y}=\underset{\sim}{r}|Z_1=l_1 m_1+1,\ Z_2=0)=P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{Z}=\underset{\sim}{0})$ and using (1.4) we have (by an argument similar to that employed by PANARETOS [4] in the univariate case)

$$\frac{g_{r_1+l_1 m_1, r_2}}{c_{r_1+l_1 m_1, r_2}} = \frac{g_{0, r_2}}{c_{0, r_2}}\, \theta_1^{r_1+l_1 m_1+1} \tag{1.8}$$

$r_i=0, 1, \dots, m_i$, $i=1, 2$ for some $\theta_1>0$ and for every $l_1=0, 1, \dots, k_1-1$.

Similarly from

$$P(\underset{\sim}{Y} = \underset{\sim}{r}|Z_1 = 0, \quad Z_2 = l_2 m_2+1) = P(\underset{\sim}{Y} = \underset{\sim}{r}|\underset{\sim}{Z} = \underset{\sim}{0})$$

we find

$$\frac{g_{r_1, r_2+l_2 m_2+1}}{c_{r_1, r_2+l_2 m_2+1}} = \frac{g_{r_1, 0}}{c_{r_1, 0}}\, \theta_2^{r_2+l_2 m_2+1} \tag{1.9}$$

$r_i=0, 1, \dots, m_i$; $i=1, 2$ for some $\theta_2>0$ and for every $l_2=0, 1, \dots, k_2-1$.

Moreover, from $P(\underset{\sim}{Y}=\underset{\sim}{r}|Z_1=l_1 m_1+1,\ Z_2=l_2 m_2+1)=P(\underset{\sim}{Y}=\underset{\sim}{r}|Z_1=l_1 m_1+1,$ $Z_2=0)$, for fixed $0 \leqq l_1 \leqq k_1-1$ and for every $l_2=0, 1, \dots, k_2-1$ we find

$$\frac{g_{r_1+l_1 m_1+1, r_2+l_2 m_2+1}}{c_{r_1+l_1 m_1+1, r_2+l_2 m_2+1}} = \frac{g_{r_1+l_1 m_1+1, 0}}{c_{r_1+l_1 m_1+1, 0}}\, \theta_3^{r_2+l_2 m_2+1} \tag{1.10}$$

$$r_i = 0, 1, \dots, m_i; \quad i = 1, 2, \quad \text{and some} \quad \theta_3 > 0.$$

It can be checked easily that $\theta_3=\theta_2$. On the other hand, l_1 was fixed but arbitrary. Hence (1.10) is valid for every

$$l_1, l_2\colon l_i = 0, 1, \dots, k_i-1; \quad i = 1, 2.$$

Combining (1.10) and (1.8) (for $r_2=0$) yields

$$\frac{g_{n_1, n_2}}{c_{n_1, n_2}} = \frac{g_{0,0}}{c_{0,0}}\, \theta_1^{n_1} \theta_2^{n_2}$$

for some $\theta_1, \theta_2 > 0$ and for all

$$n_i = 0, 1, \ldots, k_i m_i + 1, \quad 1 \leqq k_i \leqq \left[\frac{N_i - 1}{m_i}\right], \quad N_i > 0, \quad i = 1, 2.$$

This completes the proof.

The theorem that has just been established extends to multivariate discrete distributions with finite support a similar existing result concerning multivariate discrete distributions with infinite support (PANATEROS [2], [5]).

3. Some Applications

As a consequence of Theorem 2 the following characterizations of the multiple binomial and the multiple hypergeometric distributions can be established.

Corollary 1. *(Characterization of the Multiple Binomial Distribution.)*

Let $(\underset{\sim}{X}, \underset{\sim}{Y})$ *be a random vector with* $\underset{\sim}{X} = (X_1, X_2, \ldots, X_s)$ *and* $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_s)$ *where* X_i, Y_i, $i = 1, 2, \ldots, s$ *are non-negative, integer-valued random variables with* $\underset{\sim}{X} \geqq \underset{\sim}{Y}$, *and* $P(\underset{\sim}{X} = \underset{\sim}{0}) < 1$. *Suppose that the distribution of* $\underset{\sim}{Y} \mid (\underset{\sim}{X} = \underset{\sim}{n})$ *is multiple hypergeometric, i.e. that*

$$P(\underset{\sim}{Y} = \underset{\sim}{r} \mid \underset{\sim}{X} = \underset{\sim}{n}) = \prod_{i=1}^{s} \frac{\binom{m_i}{r_i}\binom{N_i - m_i}{n_i - r_i}}{\binom{N_i}{n_i}} \tag{3.1}$$

$$r_i, n_i, m_i, N_i > 0, \quad r_i \leqq n_i; \quad i = 1, 2, \ldots, s.$$

Then the condition (1.5) *for* $k_i = \left[\frac{N_i - 1}{m_i}\right]$, $i = 1, 2, \ldots, s$ *is necessary and sufficient for* $\underset{\sim}{X}$ *to follow a multiple binomial distribution, i.e.,*

$$P(\underset{\sim}{X} = \underset{\sim}{n}) = \prod_{i=1}^{s} \binom{N_i}{n_i} p_i^{n_i} q_i^{N_i - n_i} \tag{3.2}$$

$$n_i = 0, 1, \ldots, N_i, \ N_i > 0, \quad \textit{for some} \quad 0 < p_i < 1, \ q_i = 1 - p_i; \quad i = 1, 2, \ldots s.$$

PROOF. Consider the sequences

$$a_{\underset{\sim}{r}} = \prod_{i=1}^{s} \binom{m_i}{r_i} \quad \text{and} \quad b_{\underset{\sim}{n}} = \prod_{i=1}^{s} \binom{N_i - m_i}{n_i} \tag{3.3}$$

$$r_i = 0, 1, \ldots, m_i; \quad n_i = 0, 1, \ldots, N_i, \quad i = 1, 2, \ldots, s.$$

It follows easily that the sequence $\{c_{\underset{\sim}{n}}\}$ defined in Theorem 2 in this particular case takes the form

$$c_{\underset{\sim}{n}} = \prod_{i=1}^{s} \binom{N_i}{n_i}, \quad n_i = 0, 1, \ldots, N_i; \ i = 1, 2, \ldots, s. \tag{3.4}$$

Clearly, $a_{\underset{\sim}{n}}, b_{\underset{\sim}{n}}, c_{\underset{\sim}{n}}$ as given by (3.3) and (3.4) can be used to express (3.1) in the form $a_{\underset{\sim}{r}} b_{\underset{\sim}{n} - \underset{\sim}{r}} / c_{\underset{\sim}{n}}$. Moreover they meet all the conditions set by Theorem 2. Consequently, as a result of Theorem 2 we have that condition (1.5), for $k_i=\left[\frac{N_i-1}{m_i}\right]$, $i=1, 2, \ldots$ $\ldots, s$ holds iff $g_{\underset{\sim}{n}}$ satisfies (1.6). A direct substitution then gives that $g_{\underset{\sim}{n}}$ is multiple binomial of the form (3.2) with $p_i=\theta_i/(1+\theta_i)$ $i=1, 2, \ldots, s$.

Other characterization problems of this or similar nature where the multiple binomial distribution was involved have been studied by, among others, SRIVASTAVA and SRIVASTAVE [11], TALWALKER [12], SHANBHAG [9] ACZÉL [1], XEKALAKI [13] and PANARETOS and XEKALAKI [7].

Corollary 2. *(Characterization of the Multiple Hypergeometric Distribution). Let* $(\underset{\sim}{X}, \underset{\sim}{Y})$ *be a random vector as in Corollary* 1. *Moreover, assume that the conditional distribution of* $\underset{\sim}{Y}$ *on* $\underset{\sim}{X}=\underset{\sim}{n}$ *is of the structural form* (1.4). *Let* $\underset{\sim}{X}$ *follow a multiple binomial distribution as in* (3.2). *Then condition* (1.5) *for* $k_i=\left[\frac{N_i-1}{m_i}\right]$, $i=1, 2, \ldots, s$ *holds iff* $\underset{\sim}{Y}|(\underset{\sim}{X}=\underset{\sim}{n})$ *follows a multiple hypergeometric distribution as in* (3.1).

To prove the above corollary we need first to establish that the multiple binomial distribution is uniquely decomposable into multiple binomials. This is done in the form of the following lemma.

Lemma. *Let* $G_1(\underset{\sim}{t})$ *and* $G_2(\underset{\sim}{t})$ *be the probability generating functions* (p.g.f.'s) *of two independent random vectors* $\underset{\sim}{X}$ *and* $\underset{\sim}{Y}$, *and let* $G(\underset{\sim}{t})$ *denote their convolution. (i.e.,* $G(\underset{\sim}{t})=G_1(\underset{\sim}{t})G_2(\underset{\sim}{t})$). *Assume that* $G(\underset{\sim}{t})$ *is the p.g.f. of a multiple binomial distribution with probability distribution as in* (3.2). $\left(\text{i.e. } G(\underset{\sim}{t})=\prod_{i=1}^{s}(q_i+p_it_i)^{N_i}\right)$. *Then* $G_1(\underset{\sim}{t})$ *and* $G_2(\underset{\sim}{t})$ *are also p.g.f.'s of multiple binomial distributions with the same* $\underset{\sim}{p}$ *i.e.,* $G_1(\underset{\sim}{t})=\prod_{i=1}^{s}(q_i+p_it_i)^{m_i}$ *and* $G_2(\underset{\sim}{t})=\prod_{i=1}^{s}(q_i+p_it_i)^{n_i}$ *with* $\sum_{i=1}^{s}(m_i+n_i)=\sum_{i=1}^{s}N_i$.

PROOF. We give a proof for $s=2$. Then the general case follows easily.

We are given that

$$G_1(t_1, t_2)G_2(t_1, t_2)=(q_1+p_1t_1)^{N_1}(q_2+p_2t_2)^{N_2}. \tag{3.5}$$

Dividing (3.5) be its value at $t_1=1$ gives

$$\frac{G_1(t_1, t_2)}{G_1(1, t_2)}\frac{G_2(t_1, t_2)}{G_2(1, t_2)}=(q_1+p_1t_1)^{N_1}. \tag{3.6}$$

Clearly, $\frac{G_1(t_1, t_2)}{G_1(1, t_2)}$ and $\frac{G_2(t_1, t_2)}{G_2(1, t_2)}$ are valid p.g.f.'s in t_1. Therefore, since the lemma is valid in the univariate case (see e.g. RAMACHANDRAN [8]), (3.6) is equivalent to

$$\frac{G_1(t_1, t_2)}{G_1(1, t_2)}=(q_1+p_1t_1)^{m_1}, \quad \frac{G_2(t_1, t_2)}{G_2(1, t_2)}=(q_1+p_1t_1)^{n_1} \tag{3.7}$$

$$m_1+n_1=N_1.$$

If we use the same technique for $t_2=1$ we will find that

$$\frac{G_1(t_1, t_2)}{G_1(t_1, 1)} = (q_2+p_2t_2)^{m_2}, \quad \frac{G_1(t_1, t_2)}{G_2(t_1, 1)} = (q_2+p_2t_2)^{n_2} \tag{3.8}$$

$$m_2+n_2 = N_2$$

Combining the first relation of (3.7) for $t_2=1$ with the first of the relations in (3.8) gives

$$G_1(t_1, t_2) = (q_1+p_1t_1)^{m_1}(q_2+p_2t_2)^{m_2}. \tag{3.9}$$

Similarly combining the second of the relations in (3.7) for $t_1=1$ with the second of the relations in (3.8) yields

$$G_2(t_1, t_2) = (q_1+p_1t_1)^{n_1}(q_2+p_2t_2)^{n_2} \tag{3.10}$$

where $m_1+m_2+n_1+n_2=N_1+N_2$. Hence (3.9) and (3.10) complete the proof of the lemma.

PROOF OF COROLLARY 2. First of all it can be checked easily that "necessity" is a side result of Corollary 1. To prove the "sufficient" part we first observe that since all the assumptions for theorem 2 hold we have that the conditions (1.5) for $k_i=\left[\frac{N_i-1}{m_i}\right]$, $i=1, 2, \ldots, s$ hold iff

$$c_{\underset{\sim}{n}} = c_{\underset{\sim}{0}} \prod_{i=1}^{s} \binom{N_i}{n_i}\left(\frac{p_i}{q_i\theta_i}\right)^{n_i}$$

which gives

$$\frac{c_{\underset{\sim}{n}}}{\sum_{\underset{\sim}{n}} c_{\underset{\sim}{n}}} = \prod_{i=1}^{s} \binom{N_i}{n_i} \pi_i^{n_i}(1-\pi_i)^{N_i-n_i}$$

where $\pi_i = \dfrac{p_i}{q_i\theta_i+p_i}$, $i = 1, 2, \ldots, s$.

So, $c_{\underset{\sim}{n}}^{\infty}$ multiple binomial distribution and therefore, according to the lemma and the definition of $c_{\underset{\sim}{n}}$, the same holds true for $a_{\underset{\sim}{r}}$ and $b_{\underset{\sim}{n}-\underset{\sim}{r}}$. The result follows easily if we take into account the fact that $P(\underset{\sim}{Y}=\underset{\sim}{r}|\underset{\sim}{X}=\underset{\sim}{n})$ was assumed to be of the structural form (1.4) and if we substitute in it $a_{\underset{\sim}{r}}, b_{\underset{\sim}{n}}, c_{\underset{\sim}{n}}$.

Note. The reader may observe that here, as in the univariate case, the number of conditions needed to characterize the multiple binomial, or the multiple hypergeometric distribution, depends on m_i and N_i, $i=1, 2, \ldots, s$. The smaller m_i is in relation to N_i the fewer the number of conditions required is.

References

[1] J. Aczél, Some Recent Applications of Functional Equations and Inequalities to Characterizations of Probability Distributions, Combinatorics, Information Theory and Mathematical Economics. In *Statistical Distributions in Scientific Work, Vol* 3, *Patil, Kotz* & *Ord (eds.)* (1975), 321—337. *D. Reidel Pub. Co. Holland.*

[2] J. Panaretos, A Characterization of a General Class of Multivariate Discrete Distributions. In Analytic Function Methods in Probability Theory. (*Colloquia Mathematica, Societatis János Bolyai,* No. **21**.) *Gyires B. (Ed.) North Holland.* (1980), 243—252.

[3] J. Panaretos, On Characterizing Some Discrete Distributions Using an Extension of the Rao—Rubin Theorem. *Sankhyā A,* **44** (3) (1982), 415—422.

[4] J. Panaretos, On the Joint Distribution of Two Discrete Random Variables. *Ann. Inst. Statist. Math.,* **33**, *Part A,* (1981), 191—198.

[5] J. Panaretos, A Characterization of the Negative Multinomial Distribution. In *Statistical Distributions in Scientific Work. Vol.* 4 *(Models, Structures and Characterizations). C. Taille, G. P. Patil and B. Baldessari (Eds.), D. Reidel Publ. Co. New York* (1981), 331—341.

[6] J. Panaretos, On a Structural Property of Finite Distributions. *J. Roy. Stat. Soc., Series B,* **44** (2), (1982), 209—211.

[7] J. Panaretos, and E. Xekalaki, A Characteristic Property of Certain Discrete Distributions. In Analytic Function Methods in Probability Theory. (*Colloquia Mathematica Sociatatis János Bolyai,* No. **21**), *Gyires B. (Ed.) North Holland* (1980), 253—267.

[8] B. Ramachandran, Advanced Theory of Characteristic Functions. Statistical Publishing Society Calcutta (1967).

[9] D. N. Shanbhag, An Elementary Proof for the Rao—Rubin Characterization of the Poisson Distribution. *J. Appl. Prob.* **11** (1974), 211—215.

[10] D. N. Shanbhag, An Extension of the Rao—Rubin Characterization of the Poisson Distribution. *J. App. Prob.,* **14** (3) (1977), 640—646.

[11] R. C. Srivastava and A. B. L. Srivastava, On a Characterization of Poisson Distributions *J. Appl. Prob.,* **7** (1970), 497—501.

[12] S. Talwalker, A Characterization of the Double Poisson Distribution *Sankhyā:* A, **32** (1970), 265—270.

[13] E. Xekalaki, On Characterizing the Bivariate Poisson Binomial and Negative Binomial Distributions. In Analytic Function Methods in Probability Theory. (*Colloquia Mathematica Sociatatis János Bolyai, No.* **21**.) *Gyires B. (Ed.) North Holland* (1980), 369—379.

UNIVERSITY OF CRETE
HERAKLIO, CRETE, GREECE.

(Received March 25, 1981)