Ann. Inst. Statist. Math.
33 (1981), Part A, 191-198

ON THE JOINT DISTRIBUTION OF TWO DISCRETE RANDOM VARIABLES

JOHN PANARETOS

(Received May 12, 1980; revised Apr. 13, 1981)

1. Introduction

Properties of discrete distributions based on the conditional distribution of a random variable (r.v.) Y for a given value of a r.v. X, where $X \geqq Y$, have attracted always attention. In most cases, (see Moran [2], Chatterji [1], Patil and Seshadri [5]) the basic assumption was that of independence between Y and $X-Y$. A very general result along this line was the one of Patil and Seshadri [5]. They essentially showed that when Y and $X-Y$ are independent and the distribution $s(r|n)$ of $Y|(X=n)$ is of the form $a_r b_{n-r}/c_n$ (where a_n, b_n, $n=0, 1, \cdots$ are sequences of non-negative real numbers whose convolution is c_n), then the distributions of Y and $X-Y$ are of a power series form. As corollaries of their result Patil and Seshadri showed that Y and $X-Y$ are Poisson r.v.'s when $s(r|n)$ is binomial, negative binomial r.v.'s when $s(r|n)$ is negative hypergeometric, and binomial r.v.'s when $s(r|n)$ is hypergeometric.

Shanbhag [7] adopted the same form for the distribution of $Y|(X=n)$ but his result is more general, since he replaced the assumption of full independence between Y and $X-Y$ by an assumption of partial independence. Roughly speaking Shanbhag proved that if $Y|(X=n)$ is

AMS 1979 subject classification: 60E05.
Key words and phrases: Conditional distribution, power series distribution, binomial distribution, characterization.

of the form $a_r b_{n-r}/c_n$ with $a_n>0$, $n=0, 1, \cdots$ and $b_0, b_1>0$; $b_n \geqq 0$, $n\geqq 2$, then $P(Y=r)=P(Y=r|X=Y)$ if and only if (iff) X has a power series distribution. (The condition $P(Y=r)=P(Y=r|X=Y)$ was employed earlier by Rao and Rubin [6] to characterize the Poisson distribution when $Y|(X=n)$ was binomial (n, p) with p fixed and independent of n. Since then this condition has become known as Rao-Rubin condition (R-R condition)).

In a forthcoming paper Panaretos [3] extends Shanbhag's result to characterize left-truncated distributions (distributions of r.v.'s taking values $\geqq k$, $k=0, 1, \cdots$) by assuming that $a_n>0$ for $n\geqq k$ only. The characterizing condition in this case is $P(Y=r|Y\geqq k)=P(Y=r|X=Y)$. Using Shanbhag's theorem and Panaretos's extension characterizations can be obtained for the Poisson distribution ($Y|(X=n)\sim$binomial), the negative binomial distribution ($Y|(X=n)\sim$negative hypergeometric), the left-truncated Poisson distribution ($Y|(X=n)\sim$binomial), the left-truncated negative binomial distribution ($Y|(X=n)\sim$negative hypergeometric), the convolution of a Poisson with a left-truncated Poisson ($Y|(X=n)\sim$left-truncated binomial), and the convolution of a negative binomial with a left-truncated negative binomial ($Y|(X=n)\sim$left-truncated negative hypergeometric). In all the above mentioned results it is apparent that the unique form of the bivariate distribution of (X, Y) can also be derived once the distribution of X is established.

The results of Shanbhag [7] and Panaretos [3] constitute a substantial generalization of Patil and Seshadri's characterization in the discrete case. However, characterizations of distributions with finite range cannot be obtained directly from them. Consider for example the case where

$$s(r|n)=\binom{m}{r}\binom{N-m}{n-r}\Big/\binom{N}{n} \qquad r\leqq n,\ m<N,\ m, n, N>0 .$$

Clearly, $s(r|n)$ can be expressed in the form $a_r b_{n-r}/c_n$ with $a_n=\binom{m}{n}$, $b_n=\binom{N-m}{n}$. However, the results of Shanbhag [7] and Panaretos [3] cannot be applied since $a_n=0$ for $n>m$.

This implies that in such cases the R-R condition, as it stands, is not adequate to replace the assumption of independence considered by Patil and Seshadri. A more stringent condition is therefore required. A general result providing an answer to this problem is stated and proved in the next section (Section 2). Then, in Section 3 some interesting examples are given.

2. The main result

THEOREM 2.1. *Let N and m be positive integers such that $N>m$, and let k_0 be the integral part of $(N-1)/m$ (i.e. $k_0=[(N-1)/m]$). Let $\{(a_n, b_n);\ n=0, 1, \cdots\}$ be a sequence of non-negative real vectors such that $a_n>0$, $n=0, 1, \cdots, m$; $b_j>0$, $j=0, 1, m+1, 2m+1, \cdots, (k-1)m+1$, for some integer k, $0<k\leqq k_0$. For $n=0, 1, \cdots, N$, put $c_n=\sum_{r=0}^{n} a_r b_{n-r}$. Consider a random vector (X, Y) of non-negative, integer-valued components such that $\mathrm{P}(X=n)=P_n$, $n=0, 1, \cdots, N$ with $P_0<1$. Suppose that*

$$\text{(2.1)} \qquad \mathrm{P}(Y=r\,|\,X=n)=\frac{a_r b_{n-r}}{c_n}$$

$$\text{whenever } P_n>0,\ r=0, 1, \cdots, n;\ n=0, 1, \cdots, N.$$

If $0<k\leqq k_0$, then the k relations

$$\text{(2.2)} \qquad \mathrm{P}(Y=r\,|\,X=Y)=\mathrm{P}(Y=r\,|\,X=Y+(j-1)m+1)\,, \quad j=1, 2, \cdots, k$$

hold iff

$$\text{(2.3)} \qquad P_n=P_0\frac{c_n}{c_0}\theta^n\,, \qquad \text{for some } \theta>0,\ n=0, 1, \cdots, km+1\,.$$

PROOF. "If" part. From (2.1) and (2.3) we have, for $0\leqq j\leqq N$

$$\mathrm{P}(Y=r\,|\,X=Y+j)=\frac{\mathrm{P}(Y=r,\ X=r+j)}{\mathrm{P}(X-Y=j)}=\frac{(a_r b_j/c_{r+j})P_{r+j}}{\mathrm{P}(X-Y=j)}$$

$$=\frac{b_j a_r}{\mathrm{P}(X-Y=j)}\frac{P_0}{c_0}\theta^{r+j} \qquad r=0, 1, \cdots\,.$$

Hence

$$\text{(2.4)} \qquad \mathrm{P}(Y=r\,|\,X=Y+j)=\frac{a_r\theta^r}{\phi(j,\theta)} \qquad r=0, 1, \cdots\,.$$

But $\sum_r \mathrm{P}(Y=r\,|\,X=Y+j)=1$. So,

$$\text{(2.5)} \qquad \phi(j,\theta)=\sum_r a_r\theta^r=A(\theta)$$

and therefore

$$\text{(2.6)} \qquad \mathrm{P}(Y=r\,|\,X=Y+j)=\frac{a_r\theta^r}{A(\theta)}\,, \qquad r=0, 1, \cdots;\ 0\leqq j\leqq N\,.$$

The fact that the right-hand side of (2.6) does not depend on j implies that $P(Y=r|X-Y=j)$ are independent of j for $0\leqq j\leqq N$ and hence these are all equal for a given r. So, (2.2) is established.

The 'only if', part can be proved by an inductive argument. For $k=1$ we are given that

$$P(Y=r|X=Y+1)=P(Y=r|X=Y) .$$

By making use of (2.1) we obtain

$$\frac{P_{r+1}\, a_r b_1/c_{r+1}}{P(X=Y+1)}=\frac{P_r\, a_r b_0/c_r}{P(X=Y)}, \qquad r=0, 1, \cdots, N-1$$

and since $a_r>0$, $r=0, 1, \cdots, m$

$$\frac{P_{r+1}}{c_{r+1}}=\frac{P_r}{c_r}\theta_{0,1}, \qquad r=0, 1, \cdots, m \tag{2.7}$$

where

$$\theta_{0,1}=\frac{b_0}{b_1}\frac{P(X=Y+1)}{P(X=Y)} \qquad \text{is a constant}. \tag{2.8}$$

Consequently

$$\frac{P_{r+1}}{c_{r+1}}=\frac{P_0}{c_0}\theta_{0,1}^{r+1} \qquad r=0, 1, \cdots, m . \tag{2.9}$$

Hence the result is true for $k=1$.

Assume now that the result is valid for $k=i$, $1\leqq i\leqq k_0$ i.e. assume that the conditions

$$\begin{aligned} P(Y=r|X=Y)&=P(Y=r|X=Y+1)\\ &=P(Y=r|X=Y+m+1)=\cdots\\ &=P(Y=r|X=Y+(i-1)m+1) \end{aligned}$$

imply that

$$\frac{P_{r+(i-1)m+1}}{c_{r+(i-1)m+1}}=\frac{P_0}{c_0}\theta_{0,1}^{r+(i-1)m+1} \qquad r=0, 1, \cdots, m . \tag{2.10}$$

We will show that the result is also valid for $k=i+1$. From

$$P(Y=r|X=Y+im+1)=P(Y=r|X=Y+(i-1)m+1)$$

we have

$$\frac{P_{r+im+1}}{c_{r+im+1}}=\frac{P_{r+(i-1)m+1}}{c_{r+(i-1)m+1}}\theta_{(i-1)m+1,im+1} \qquad r=0, 1, \cdots, m$$

with $\theta_{(i-1)m+1,im+1}=\dfrac{b_{(i-1)m+1}}{b_{im+1}}\dfrac{P(X=Y+im+1)}{P(X=Y+(i-1)m+1)}$, a constant. Hence, taking into consideration (2.10),

$$\frac{P_{r+im+1}}{c_{r+im+1}}=\frac{P_0}{c_0}\theta_{0,1}^{r+(i-1)m+1}\theta_{(i-1)m+1,\,im+1} \qquad r=0, 1, \cdots, m\,. \tag{2.11}$$

Relations (2.10) and (2.11) yield

$$\theta_{(i-1)m+1,\,im+1}=\theta_{0,1}^{m}\,.$$

Hence, finally

$$\frac{P_{r+im+1}}{c_{r+im+1}}=\frac{P_0}{c_0}\theta_{0,1}^{r+im+1}\,, \qquad r=0, 1, \cdots, m\,. \tag{2.12}$$

So, the statement is also true for $k=i+1$. This completes the induction argument. Consequently, we have that (2.2) implies (2.3), and the theorem is established.

Note 2.1. Evidently, once the form of the distribution of X is unique the same is true for the joint distribution of (X, Y).

Note 2.2. The theorem just proved is a variant of Shanbhag's [7] theorem and a generalization of the discrete case of Patil and Seshadri's [5] result. (Patil and Seshadri are making use of the fact that Y and $X-Y$ are independent. This, for finite distributions is equivalent to the $N+1$ relations $\mathrm{P}(Y=r)=\mathrm{P}(Y=r\,|\,X=Y+l)$ for all $l=0, 1, \cdots, N-r$; $r=0, 1, \cdots, N$. By (2.1) only $k<N$ relations are needed. Moreover, these k relations do not involve the unconditional distribution of Y.) It also provides an answer to the problem of characterizing discrete distributions with finite support as stated in the previous section. In fact, Theorem 2.1 tells us more. It indicates that if one raises the number k by 1 in (2.2) then one can determine another m probabilities from the probability distribution P_n $(n=0, 1, \cdots, N)$. This is evident in the course of the proof of the theorem. To characterize completely a distribution with finite range one should appeal to the extreme case of the theorem where $k=k_0$. (In this case it is clear that all the N frequencies of P_n are characterized in terms of the k_0 conditions (2.2).) This implies that to characterize a finite distribution, the number of R-R type conditions required depends on the parameters (N, m).

Note 2.3. It is interesting to observe that if the distribution of Y given X is of the form (2.1) and the distribution of X satisfies (2.3) for $n=0, 1, \cdots, N$ then Y and $X-Y$ are independent.

3. Some Applications

Let us consider a random vector (X, Y) of non-negative integer valued components such that $\mathrm{P}(X=n)=P_n$, $n=0, 1, \cdots, N$ with $P_0<1$.

Then, as a direct consequence of Theorem 2.1 the following corollaries can be established.

COROLLARY 3.1 (*Characterization of the Binomial Distribution*). *Suppose that*

$$\text{(3.1)} \qquad \mathrm{P}(Y=r|X=n)=\binom{m}{r}\binom{N-m}{n-r}\Big/\binom{N}{n}\,, \qquad r\leqq n,\ m<N,\ m, n, N>0$$

i.e. $s(r|n)\sim$ *hypergeometric* (m, n, N). *Then*

$$\mathrm{P}(Y=r\,|\,X=Y)=\mathrm{P}(Y=r\,|\,X=Y+(j-1)m+1)\,, \qquad j=1, 2, \cdots, k_0$$

with

$$\text{(3.2)} \qquad k_0=\left[\frac{N-1}{m}\right]$$

iff

$$\text{(3.3)} \qquad P_n=\binom{N}{n}p^n q^{N-n}\,; \qquad 0<p<1,\ q=1-p,\ N>0,\ n=0, 1, \cdots, N.$$

PROOF. Define

$$a_n=\binom{m}{n}\,, \qquad b_n=\binom{N-m}{n} \qquad n=0, 1, \cdots .$$

Then $c_n=\sum_{r=0}^{n} a_r b_{n-r}=\binom{N}{n}$, $n=0, 1, \cdots, N$.

These sequences can be used to express (3.1) in the form $a_r b_{n-r}/c_n$ and also satisfy the requirements of Theorem 2.1. Consequently, applying the result of the theorem we find that (3.2) holds iff P_n is binomial as in (3.3). (A reference to the problem solved by this corollary is made in the paper by Shanbhag and Taillie [9]).

Note 3.1. Patil and Ratnaparkhi [4] have observed that if the conditional distribution of Y given X is hypergeometric (m, n, N) as in (3.1) and $X\sim$binomial $(n; N, p)$ then the R-R condition holds. It can now be seen that their remark is a side result of the 'if' part of Corollary 3.1. In the same paper the authors raised the question as to whether the R-R condition implies that $X\sim$binomial $(n; N, p)$ when $s(r|n)\sim$hypergeometric (m, n, N). Shanbhag and Panaretos [8] by means of a counter example showed that this is not necessarily the case. By means of Corollary 3.1 we are now in a position to point out that in general $[(N-1)/m]$ conditions of the R-R type are required for X to be binomial $(n; N, p)$. There is however, a special case in which only one condition of the R-R type is adequate to characterize the binomial distribution (with $Y|(X=n)\sim$hypergeometric (m, n, N)). This case arises when m takes certain values in relation to N. Formally this case can be stated as follows.

COROLLARY 3.2 (*Characterization of the Binomial distribution using a Rao-Rubin type condition*). *Suppose that*

$$\mathrm{P}(Y=r\,|\,X=n)=\binom{m}{r}\binom{N-m}{n-r}\Big/\binom{N}{n}, \quad r\leqq n\,; \quad \frac{N-1}{2}<m<N\,; \; n, N>0\,.$$

i.e. $s(r\,|\,n)\sim$ *hypergeometric* (m, n, N) $((N-1)/2<m<N)$. *Then*

$$\mathrm{P}(Y=r\,|\,X=Y)=\mathrm{P}(Y=r\,|\,X=Y+1)$$

iff $X\sim$*binomial* $(n; N, p)$ *as in* (3.3).

PROOF. From the previous corollary we have that $X\sim$binomial $(n; N, p)$ iff (3.2) holds. But because of the additional restriction imposed on m, we see that we should have $k_0<2$. Since k_0 is a positive integer, the result follows.

Remark. It is evident from our previous comments that as m increases the number k_0 of the R-R type conditions required to characterize the binomial distribution decreases. This number takes its minimum value ($k_0=1$) when m exceeds $(N-1)/2$.

Acknowledgements

I would like to thank the referee for his comments that helped to improve the presentation of Theorem 2.1.

TRINITY COLLEGE, DUBLIN*

REFERENCES

[1] Chatterji, S. K. (1963). Some elementary characterizations of the Poisson distribution, *Amer. Math. Monthly*, **70**, 958-964.

[2] Moran, P. A. P. (1952). A characteristic property of the Poisson distribution, *Proc. Camb. Phil. Soc.*, **48**, 206-207.

[3] Panaretos, J. (1979). On characterizing some discrete distributions using an extension of the Rao-Rubin theorem, to appear in *Sankhyā*.

[4] Patil, G. P. and Ratnaparkhi, M. V. (1975). Problems of damaged random variables and related characterizations, *Statist. Distrib. in Scient. Work*, Vol. 3 (eds. G. P. Patil, S. Kotz and J. K. Ord), D. Reidal Publ. Co., Utrecht, 255-270.

* Now at University of Missouri-Columbia.

[5] Patil, G. P. and Seshadri, V. (1964). Characterization theorems for some univariate discrete distributions, *J. Roy. Statist. Soc.*, B, **26**, 286-292.

[6] Rao, C. R. and Rubin, H. (1964). On a characterization of the Poisson distribution, *Sankhyā*, A, **26**, 295-298.

[7] Shanbhag, D. N. (1977). An extension of the Rao-Rubin characterization of the Poisson distribution, *J. Appl. Prob.*, **14**, 640-646.

[8] Shanbhag, D. N. and Panaretos, J. (1979). Some results related to the Rao-Rubin characterization of the Poisson distribution, *Aust. J. Statist.*, **21**, 78-83.

[9] Shanbhag, D. N. and Taillie, C. (1980). An extension of the Patil-Taillie characterization of the Poisson distribution, to appear in *Sankhyā*.