On a Structural Property of Finite Distributions

Panaretos, John

1982

J. R. Statist. Soc. B (1982),
44, *No.* 2, *pp.* 209–211

On a Structural Property of Finite Distributions

By JOHN PANARETOS†

University of Dublin, Trinity College, Dublin, Ireland

[Received June 1980. Revised June 1981]

1. INTRODUCTION

THE problem of deriving the distribution of the sum $X = Y+Z$ of two random variables Y and Z (or equivalently the distributions of Y and Z) when the conditional distribution $s(y\mid x)$ of one of them, say Y, on X is known has received a lot of attention in distribution theory. Research in this area has been mainly concerned with specifying the sort of additional information one should have to be able to solve this problem. In most cases this additional information takes the form of independence between Y and Z. In the case of discrete random variables important contributions have been made by, among others, Moran (1952), and Patil and Seshadri (1964), for specific forms of $s(y\mid x)$. From time to time, attempts have been made to relax the assumption of independence between Y and Z. One of the pioneer papers in this direction was that of Rao and Rubin (1964). They replaced the assumption of independence by the condition $P(Y = y) = P(Y = y\mid Z = 0)$ to characterize the Poisson distribution as the distribution of X when $s(y\mid x)$ was binomial with parameters p, x; p in $(0, 1)$, fixed and independent of x. (Shanbhag and Panaretos, 1979, among other things, pointed out the similarities and limitations of Rao and Rubin's paper in connection with Moran's work). Shanbhag (1977) and Panaretos (1979) effectively solved the problem of determining the class of distributions for X using Rao and Rubin's condition under the weaker assumption that $s(y\mid x)$ was of the form $a_y b_{x-y}/c_x$ where a_n, b_n are non-negative sequences satisfying certain conditions and c_n is the convolution of a_n and b_n.

A more general problem is to express the distribution of X in terms of the distribution $s(y\mid x)$ for a general $s(y\mid x)$ (i.e. without requiring $s(y\mid x)$ to have a specific structural form). Patil and Seshadri (1964) solved this problem in the discrete case for distributions with finite support under the assumption of independence between Y and Z (Patil and Seshadri, 1964, Theorem 3). They expressed the distributions of Y and Z (and hence the distribution of X) in terms of $s(y\mid x)$. An interesting question generated by Theorem 3 of Patil and Seshadri is whether the distribution of X can still be determined in terms of $s(y\mid x)$ if the assumption of indepedence of Y and Z is replaced by a weaker condition. This paper deals precisely with this problem. Specifically, in the next section we first state (in our notation) Theorem 3 of Patil and Seshadri (1964) for ease of reference and draw attention to some additional restrictions needed so that the result be always meaningful. We then state and prove our main theorem which extends Patil and Seshadri's result to non-independent Y, Z.

2. PATIL AND SESHADRI'S THEOREM AND ITS EXTENSION

Theorem 1. (Patil and Seshadri, 1964.) Let Y, Z be independent, non-degenerate, non-negative, integer-valued random variables such that $a \leqslant Y \leqslant a+k$ and $b \leqslant Z \leqslant b+m$ with

† Now at the University of Missouri–Columbia, Department of Statistics, Columbia, Mo., 65211, USA

0035–9246/82/44209$2.00

$P(Y = y)>0$, $a \leqslant y < a+k$ and $P(Z = z)>0$, $b \leqslant z < b+m$, respectively. Let $X = Y+Z$ and let $s(y \mid x)$ denote $P(Y= y \mid X = x)$. Then

$$P(Y = y) = A(y)\,\theta^y \Big/ \sum_{y=a}^{a+k} A(y)\,\theta^y \quad \text{and} \quad P(Z = z) = B(z)\,\lambda^z \Big/ \sum_{z=b}^{b+m} B(z)\,\lambda^z,$$

where

$$\theta = \frac{P(Z = b+1)}{P(Z = b)}, \quad \lambda = \frac{P(Y = a+1)}{P(Y = a)},$$

$$A(y) = \prod_{i=1}^{y-a} \{s(a+i \mid a+b+i)/s(a+i-1 \mid a+b+i)\}$$

and

$$B(z) = \prod_{j=1}^{z-b} \{s(b+j \mid a+b+j)/s(b+j-1 \mid a+b+j)\}.$$

Note that the theorem was originally stated with the assumptions that $P(Y = y)>0$ for $y = a$ only and $P(Z = z)>0$ for $z = b$ only which clearly are not adequate for $A(y)$ and $B(z)$ to be well defined. Observe also that it is not necessary for the validity of the theorem to have either $k<\infty$ or $m<\infty$.

Theorem 2. Consider a random vector (X, Y) of non-negative, integer-valued components such that $P(X = x) = g_x$, $x = 0, 1, \ldots, N$, with $g_0<1$. Denote by $[a]$ the integral part of a and let $k_0 = [(N-1)/m]$. Let $P(Y = y \mid X = x) = s(y \mid x)$, $y = 0, 1, \ldots, x$, with $\{s(y \mid x)\colon y = 0, 1, \ldots, x\}$ as distributions such that, for $y = 0, 1, \ldots, m(m \leqslant x)$, $s(y \mid y)>0$, $s(y \mid y+\rho m+1)>0$; $\rho = 0, 1, \ldots, k-1$, where k is an integer, $1 \leqslant k \leqslant k_0$. Let

$$\begin{aligned} P(Y = y \mid X = Y) &= P(Y = y \mid X = Y+1) = P(Y = y \mid X = Y+m+1) = \ldots \\ &= P(Y = y \mid X = Y+(k-1)m+1). \end{aligned} \tag{2.1}$$

Then

$$g_x = g_0\, A(x)\, \lambda_0^x, \quad x = 0, 1, \ldots, km+1,$$

where λ_0 is a positive constant and $A(x)$ is a function depending only on $s(\,.\mid x)$.

Proof. We first observe that the conditions imposed on $s(y \mid x)$ imply that $g_x>0$ for all $x = 0, 1, \ldots, N$. We also see that the system of equations (2.1), for $y = 0, 1, \ldots, m$ implies

$$\begin{aligned} g_{y+1} &= \lambda_0\, g_y \frac{s(y \mid y)}{s(y \mid y+1)}, \\ g_{y+\rho m+1} &= \lambda_0\, g_{y+(\rho-1)m+1} \frac{s(y \mid y+(\rho-1)m+1)}{s(y \mid y+\rho m+1)} \end{aligned} \tag{2.2}$$

for all $y = 0, 1, \ldots, m$ and all $\rho = 1, 2, \ldots, k-1$ where λ_i, $i = 0, 1, \ldots, k-1$, are constants given by

$$\lambda_0 = \frac{P(X = Y+1)}{P(X = Y)}, \quad \lambda_i = \frac{P(X = Y+im+1)}{P(X = Y+(i-1)m+1)}; \quad i = 1, 2, \ldots, k-1.$$

Multiplication by parts of the k equations in (2.2) yields

$$\begin{gathered} g_{y+\rho m+1} = \lambda_{(\rho)}\, g_y \frac{s(y \mid y)}{s(y \mid y+\rho m+1)}, \\ \rho = 0, 1, \ldots, k-1; \quad y = 0, 1, \ldots, m \quad \text{and} \quad \lambda_{(\rho)} = \prod_{i=0}^{\rho} \lambda_i. \end{gathered} \tag{2.3}$$

Using (2.3) we can see that

$$g_{y+\rho m+1} = g_0 \lambda_0^y \lambda_{(\rho)} \prod_{i=0}^{y} \frac{s(i\,|\,i)}{s(i\,|\,i+1)} \frac{s(y\,|\,y+1)}{s(y\,|\,y+\rho m+1)}, \qquad y = 0, 1, \ldots, m; \quad \rho = 0, 1, \ldots, k-1. \tag{2.4}$$

It can now be checked that, for $\rho = 1, 2, \ldots, k-1$,

$$\lambda_{(\rho)} = \lambda_0^m \lambda_{(\rho-1)} s(0\,|\,\rho m+1) \frac{s(m\,|\,m+1)}{s(m\,|\,\rho m+1)} \prod_{i=0}^{m} \frac{s(i\,|\,i)}{s(i\,|\,i+1)} \tag{2.5}$$

and therefore

$$\lambda_{(\rho)} = \lambda_0^{\rho m+1} \frac{s(m\,|\,1)}{s(0\,|\,1)} \left\{ \prod_{j=0}^{\rho} \frac{s(0\,|\,jm+1)}{s(m\,|\,jm+1)} \right\} \left\{ s(m\,|\,m+1) \prod_{i=0}^{m} \frac{s(i\,|\,i)}{s(i\,|\,i+1)} \right\}^{\rho} \qquad \rho = 0, 1, \ldots, k-1. \tag{2.6}$$

Combining (2.4) and (2.6) yields

$$g_{y+\rho m+1} = g_0 \lambda_0^{y+\rho m+1} A(y, \rho, m), \qquad y = 0, 1, \ldots, m; \quad \rho = 0, 1, \ldots, k-1; \quad \lambda_0 \text{ a constant} \tag{2.7}$$

and

$$A(y, \rho, m) = \frac{s(y\,|\,y+1)}{s(y\,|\,y+\rho m+1)} \left\{ \prod_{i=0}^{y} \frac{s(i\,|\,i)}{s(i\,|\,i+1)} \right\} \frac{s(m\,|\,1)}{s(0\,|\,1)} \left\{ \prod_{j=0}^{\rho} \frac{s(0\,|\,jm+1)}{s(m\,|\,jm+1)} \right\} \times \left\{ s(m\,|\,m+1) \prod_{i=0}^{m} \frac{s(i\,|\,i)}{s(i\,|\,i+1)} \right\}^{\rho}. \tag{2.8}$$

This establishes the result.

One may observe that if g_x is of the form (2.7) then (2.1) holds. Notice also that, for $k = k_0$, (2.1) completely determines g_x for all $x = 0, 1, \ldots, N$.

Remark. It is interesting to note that if, in addition to the assumptions of Theorem 2, it is also known that $s(y\,|\,x)$ is of the structural form $a_y b_{x-y}/c_x$ studied by Shanbhag (1977) and Panaretos (1979) then the relation (2.7) reduces to $g_x = g_0(c_x/c_0)\lambda^x$, $x = 0, 1, \ldots, km+1$, and $\lambda = (b_0/b_1) P(X = Y+1)/P(X = Y)$. A similar result in this special case has been obtained already by Panaretos (1981).

ACKNOWLEDGEMENT

I would like to thank Dr D. N. Shanbhag for some constructive comments.

REFERENCES

MENON, M. V. (1966). Characterization theorems for some univariate probability distributions. *J. R. Statist. Soc.* B, **28**, 143–145.

MORAN, P. A. P. (1952). A characteristic property of the Poisson distribution. *Proc. Camb. Phil. Soc.*, **48**, 206–207.

PANARETOS, J. (1979). On characterizing some discrete distributions using an extension of the Rao–Rubin theorem. *Sankhyā* A. **44** (2), 1983 (To appear).

—— (1981) On the joint distribution of two discrete random variables. *Ann. Inst. Statist. Math.*, **33** (2), (Part A), 191–198.

PATIL, G. P. and SESHADRI, V. (1964). Characterization theorems for some univariate discrete distributions. *J. R. Statist. Soc.* B, **26**, 286–292.

RAO, C. R. and RUBIN, H. (1964). On a characterization of the Poisson distribution. *Sankhyā* A, **26**, 295–298.

SHANBHAG, D. N. (1977). An extension of the Rao–Rubin characterization of the Poisson distribution. *J. Appl. Prob.*, **14**, 640–646.

SHANBHAG, D. N. and PANARETOS, J. (1979). Some results related to the Rao–Rubin characterization of the Poisson distribution. *Austral. J. Statist.*, **21**, 78–83.