Characterizations of discrete distributions using the Rao–Rubin condition

Nickos Papadatos*

Department of Mathematics, Section of Statistics and O.R., University of Athens, Panepistemiopolis, 157 84 Athens, Greece

Received 10 October 2002; received in revised form 16 December 2003; accepted 7 February 2005
Available online 22 March 2005

Abstract

Consider the multivariate splitting model \( N = N_1 + \cdots + N_k \), where \( N_1, \ldots, N_k, k \geq 3 \), are arbitrary (not necessarily independent) random variables (r.v.’s) taking values in \( \mathbb{N} = \{0, 1, \ldots\} \), and assume that the Rao–Rubin condition is satisfied for \( N_1 \) and \( N_2 \). Also assume that the conditional distribution of the vector \( (N_1, \ldots, N_k) \) given \( N \) is a convolution type. Characterizations related to this model (with \( k = 2 \)) was first considered by Shanbhag (1977. J. Appl. Probab. 14, 640–646), as an extension of the binomial damage model established by Rao and Rubin (1964. Sankhyā Ser. A 26, 295–298), and was extended to any \( k \geq 3 \) by Rao and Srivastava (1979. Sankhyā Ser. A 41, 124–128).

In the present paper we provide an alternative set of conditions, under which the distribution of \( N \) is characterized, and we apply the result to some discrete distributions.

\( \odot \) 2005 Elsevier B.V. All rights reserved.

MSC: Primary 62E10

Keywords: Rao–Rubin condition; Multivariate splitting model; Poisson; Negative binomial

1. Introduction

Moran’s (1952) characterization of Poisson distribution states that if \( N_1 \) and \( N_2 \) are non-degenerate independent random variables (r.v.’s) taking non-negative integral values...
and if the conditional distribution of \( N_1 \) given \( \{N_1 + N_2 = n\} \) is binomial with index parameter \( n \in \mathbb{N} = \{0, 1, \ldots\} \) and success probability \( p_n \in [0, 1] \) for all \( n \in \mathbb{N} \) for which \( \mathbb{P}[N_1 + N_2 = n] > 0 \), then the distributions of \( N_1, N_2 \) and \( N = N_1 + N_2 \) are Poisson, provided that for some \( i \in \mathbb{N} \), \( \mathbb{P}[N_1 = i] > 0 \) and \( \mathbb{P}[N_2 = i] > 0 \). Similar results were proved by Rényi (1964) and Srivastava (1971) for the Poisson process.

Rao and Rubin (1964) proved a version of Moran’s (1952) result, in which independence of \( N_1 \) and \( N_2 \) was relaxed to the so-called **Rao–Rubin condition** (RR for brevity), namely,

\[
\mathbb{P}[N_2 = n_2 | N_1 = 0] = \mathbb{P}[N_2 = n_2], \quad n_2 \in \mathbb{N}. \quad (1.1)
\]

Specifically, their main result asserts that if the distribution of \( N = N_1 + N_2 \) is not concentrated at 0 and if for all \( n \in \mathbb{N} \) with \( \mathbb{P}[N = n] > 0 \),

\[
\mathbb{P}[N_1 = n_1 | N = n] = \binom{n}{n_1} p_1^{n_1} (1 - p)^{n - n_1}, \quad n_1 = 0, \ldots, n \quad (1.2)
\]

for some fixed \( p \in (0, 1) \), then the RR condition (1.1) implies that \( N_1 \) and \( N_2 \) are independent Poisson’s with parameters \( \lambda p \) and \( \lambda (1 - p) \), respectively, for some \( \lambda > 0 \). In other words, the RR condition is equivalent to the independence of \( N_1 \) and \( N_2 \), under the binomial damage model (1.2). (According to Rao and Rubin (1964), we may view \( N_1 \) as the undamaged (observed) part, and \( N_2 = N - N_1 \) as the damaged (unobserved) part of a natural discrete random quantity \( N \), so that (1.2) presents a binomial destructive law—a damage model for \( N \).)


Shanbhag (1977) extended the Rao–Rubin characterization to a general, convolution type, bivariate model. The multivariate analogue of Shanbhag’s characterization, namely the general multivariate splitting model \( N = N_1 + \cdots + N_k, \ k \geq 3 \), was first considered by Rao and Srivastava (1979) as an extension to Shanbhag’s (1977) model. To this end, they used the following definition.

**Definition 1.1.** Let \( N_1, \ldots, N_k, k \geq 3, \) be arbitrary r.v.’s (independence is not imposed) taking values in \( \mathbb{N} \), and assume that \( N = N_1 + \cdots + N_k \) has p.m.f. \( f(n), n \in \mathbb{N} \), such that \( f(0) < 1 \). Suppose that the functions \( a_j : \mathbb{N} \to [0, \infty), j = 1, \ldots, k \), satisfy the following:

- (A1) \( a_2(n) > 0 \) for all \( n \in \mathbb{N} \),
- (A2) \( a_j(0) > 0 \) for all \( j = 1, \ldots, k \), and
- (A3) \( a_1(1) > 0 \).

Let \( c(n) = (a_1 * \cdots * a_k)(n) = \sum_{s_1 + \cdots + s_k = n} \prod_{j=1}^{k} a_j(s_j), \ n \in \mathbb{N} \), be the convolution of \( a_1, \ldots, a_k \). We say that the random vector \( (N_1, \ldots, N_k) \) is a convolution type if for every
\( n \in \mathbb{N} \) with \( f(n) > 0 \),

\[
\mathbb{P}[N_1 = n_1, \ldots, N_k = n_k | N = n] = \frac{1}{c(n)} \prod_{j=1}^{k} a_j(n_j)
\]

for all \( n_1, \ldots, n_k \in \mathbb{N} \) with \( n_1 + \cdots + n_k = n \). \((1.3)\)

Their main result, based on Shanbhag’s result related to the RR condition, is given here for easy reference.

**Theorem 1.1 (Rao–Srivastava characterization).** If the random vector \((N_1, \ldots, N_k)\), \(k \geq 3\), is a convolution type (according to the notation used in Definition 1.1), and if the random variables \(N_1\) and \(N_2\) satisfy the RR condition \((1.1)\), then the random variables \(N_1, \ldots, N_k\) are independent with p.m.f.’s of the form

\[
P[N_j = n_j] = A_j(c)a_j(n_j)c^{n_j}, \quad n_j \in \mathbb{N}, \ j = 1, \ldots, k
\]

(for some common \( c > 0 \)) if and only if the equations

\[
\sum_{n \geq s} \frac{1}{c(n)} b(n - s)x_n = A(c)c^{x}, \quad s = 0, 1, \ldots
\]

have a unique solution for \( x_0, x_1, \ldots \), where \( \{x_0, x_1, \ldots \} \) forms a probability distribution over \( \mathbb{N} \) and \( b(n) = (a_3 \ast \cdots \ast a_k)(n) = \sum_{s_3 + \cdots + s_k = n} \prod_{j=3}^{k} a_j(s_j), \ n \in \mathbb{N}, \) is the convolution of \( a_3, \ldots, a_k \).

Theorem 1.1 implies a multivariate version of Rao and Rubin’s (1964) characterization of the Poisson distribution, thus improving the results of Rényi (1970), Bol’shev (1965) and Gerber (1979) (see, also, Corollary 2.1). Clearly, it is not an obvious fact to check whether Eqs. \((1.5)\) have a unique solution among the probability distributions over \( \mathbb{N} \), and so, the only application of Theorem 1.2 that is available in the literature is the one related to Poisson distribution. The purpose of this article is to present an alternative set of conditions (which can be shown to be sufficient for \((1.5)\) to have a unique solution among the probability distributions), so that the conclusion of Theorem 1.1 holds true. This set of conditions enables us to give, as particular examples, some new characterizations.

2. Main result

**Theorem 2.1.** Let \((N_1, \ldots, N_k), k \geq 3, \) be a convolution type (according to the notation given in Definition 1.1), and suppose that \(N_1\) and \(N_2\) satisfy the RR condition \((1.1)\). Moreover, assume that there exists a sequence of real-valued functions \(g_n : [0, \infty) \rightarrow \mathbb{R}, n \in \mathbb{N}, \) such that the following three conditions hold.

(B1) The functions \(g_n(\cdot)\) are linearly independent, in the sense that for any real constants \(c_n, n \in \mathbb{N}, \) and for any finite interval \(I \subset [0, \infty)\) of positive length, the relations

\[
\sum_{n=0}^{\infty} |c_n g_n(x)| < \infty \quad \text{and} \quad \sum_{n=0}^{\infty} c_n g_n(x) = 0 \quad \text{for all} \ x \in I,
\]

imply that \(c_n = 0\) for all \( n \in \mathbb{N}. \)
(B2) If for some \( \theta > 0 \), \( \sum_n a_2(n) \theta^n < \infty \), then \( \sum_n |g_n(x)| t^n < \infty \) for all \( t \in [0, \theta] \) and \( x \geq 0 \), and the generating function \( G(t, x) = \sum_n g_n(x) t^n \) of the sequence \( g_n(\cdot) \) has the form

\[
G(t, x) = D \exp(xg(t)), \quad 0 \leq t \leq \theta, \quad x \geq 0,
\]

where \( g(t) \) does not depend on \( x \), and \( D > 0 \) is a real constant.

(B3) There exist real constants \( M > 0 \) and \( a \geq 0 \) such that

\[
\sum_{n \geq 0} g_m(x)b(n - m) = M g_n(x + a), \quad x \geq 0, \quad n \in \mathbb{N},
\]

where \( b(\cdot) \) is the convolution of \( a_3, \ldots, a_k \), as in Theorem 1.1.

Under the above conditions, there exists some constant \( c > 0 \) such that \( \sum_{n=0}^{\infty} a_j(n)c^n < \infty \), for all \( j = 1, \ldots, k \), and, moreover, \( N_1, \ldots, N_k \) are independent r.v.’s with p.m.f.’s given by (1.4), where \( A_j(c) \), \( j = 1, \ldots, k \) are the corresponding normalizers.

**Proof.** Using the results of Shanbhag (1977), as in the proof of Theorem 2(i) in Rao and Srivastava (1979), it follows that there exist constants \( c > 0 \) and \( A > 0 \) such that

\[
\sum_{n=0}^{\infty} a_1(n)c^n < \infty, \quad \sum_{n=0}^{\infty} a_2(n)c^n < \infty \quad \text{and} \quad h(n) = Ac^n, \quad n \in \mathbb{N},
\]

where

\[
h(n) = \sum_{m \geq n} \frac{f(m)}{c(m)} b(m - n), \quad n \in \mathbb{N}. \tag{2.1}
\]

Consider a sequence of functions \( g_n, n \in \mathbb{N}, \) satisfying (B1)–(B3). Since \( \sum_n a_2(n)c^n < \infty \), it follows that \( \sum_n h(n)|g_n(x)| = A\sum_n c^n |g_n(x)| \) is finite for all \( x \geq 0 \). Therefore, by (B2),

\[
\sum_n h(n)g_n(x) = AD \exp(xg(c))
\]

\[
= A \exp(-ag(c)) D \exp((x + a)g(c))
\]

\[
= A \exp(-ag(c)) G(c, x + a)
\]

\[
= A \exp(-ag(c)) \sum_n g_n(x + a)c^n. \tag{2.2}
\]

On the other hand, using (2.1), (B3) and Fubini’s Theorem, we have

\[
\sum_n h(n)g_n(x) = \sum_n g_n(x) \sum_{m \geq n} \frac{f(m)}{c(m)} b(m - n)
\]

\[
= \sum_{m} \left( \frac{f(m)}{c(m)} \right) \sum_{0 \leq n \leq m} g_n(x)b(m - n)
\]

\[
= M \sum_n \left( \frac{f(n)}{c(n)} \right) g_n(x + a). \tag{2.3}
\]

Thus, equating the RHSs of (2.2) and (2.3) we get

\[
\sum_{n=0}^{\infty} \left( \frac{f(n)}{c(n)} \right) g_n(y) = \sum_{n=0}^{\infty} B c^n g_n(y) < \infty \quad \text{for all} \quad y \geq a,
\]
where \( B = A \exp(-ag(c))/M > 0 \), and by assumption (B1) we get
\[
f(n) = Bc(n)e^n, \quad n \in \mathbb{N}.
\]
This implies that \( \sum_n c(n)e^n < \infty \), and in combination with (1.3), yields
\[
\mathbb{P}[N_1 = n_1, \ldots, N_k = n_k] = Ba_1(n_1) \cdots a_k(n_k)c^{n_1+\cdots+n_k}, \quad n_1, \ldots, n_k \in \mathbb{N},
\]
from which the desired result follows. \( \square \)

The following results are simple by-products of Theorem 2.1.

**Corollary 2.1** (Rao–Srivastava characterization of Poisson distribution). Suppose that the r.v.’s \( N_1, \ldots, N_k, k \geq 3 \), take values in \( \mathbb{N} \) and satisfy the RR condition (1.1), and assume that \( N = N_1 + \cdots + N_k \) has p.m.f. \( f(n), n \in \mathbb{N}, \) with \( f(0) < 1 \). If there exist \( p_1 > 0, p_2 > 0, p_3 \geq 0, \ldots, p_k \geq 0 \) with \( p_1 + \cdots + p_k = 1 \) such that for every \( n \in \mathbb{N} \) with \( f(n) > 0 \),
\[
\mathbb{P}[N_1 = n_1, \ldots, N_k = n_k | N = n] = n! \prod_{j=1}^k \frac{p_j^{n_j}}{n_j!}
\]
for all \( n_1, \ldots, n_k \in \mathbb{N} \) with \( n_1 + \cdots + n_k = n \),

(\( \text{where } p_j^{n_j} \text{ should be treated as 1 if } p_j = n_j = 0 \)), then there exists some \( \lambda > 0 \) such that \( N \) is Poisson with parameter \( \lambda \), and \( N_j, j = 1, \ldots, k, \) are independent Poisson with parameters \( \lambda p_j, j = 1, \ldots, k, \) respectively (in the sense that \( N_j = 0 \) a.s., whenever \( p_j = 0 \)).

**Proof.** The assumptions of Theorem 2.1 are satisfied with \( a_j(n) = p_j^n/n!, \quad j = 1, \ldots, k, \)
\( c(n) = 1/n!, b(n) = (1 - p_1 - p_2)^n/n!, \) and the linearly independent functions \( g_n(x) = x^n/n!, x \geq 0, n \in \mathbb{N}. \) \( \square \)

**Remark 2.1.** (a) The case \( p_3 = \cdots = p_k = 0 \) (i.e. \( p_1 + p_2 = 1 \)) in Corollary 2.1, leads to the classical Rao–Rubin characterization of Poisson.

(b) The characterizations of Bol’shev (1965), Gerber (1979) and Rényi (1970, p. 142) are strictly weaker than the assertion of Corollary 2.1, because they assume independence (at least) among \( N_1 \) and \( N_2 \).

**Corollary 2.2** (Characterization of negative binomial distribution). Suppose that the r.v.’s \( N_1, \ldots, N_k, k \geq 3 \), take values in \( \mathbb{N} \) and satisfy the RR condition (1.1), and assume that \( N = N_1 + \cdots + N_k \) has p.m.f. \( f(n), n \in \mathbb{N}, \) with \( f(0) < 1 \). If there exist \( r_1 > 0, r_2 > 0, r_3 \geq 0, \ldots, r_k \geq 0 \) such that for every \( n \in \mathbb{N} \) with \( f(n) > 0 \),
\[
\mathbb{P}[N_1 = n_1, \ldots, N_k = n_k | N = n] = \left( \prod_{j=1}^k \left[ \begin{array}{c} r_j \\ n_j \end{array} \right] \right) / \left[ \begin{array}{c} r \\ n \end{array} \right]
\]
for all \( n_1, \ldots, n_k \in \mathbb{N} \) with \( n_1 + \cdots + n_k = n, \)
where \( r = r_1 + \cdots + r_k > 0 \) and

\[
\begin{cases}
1, & \text{if } n = 0, \\
\frac{1}{n!} \prod_{s=0}^{n-1} (x + s) & \text{if } n \geq 1
\end{cases}
\]

for all \( x \in \mathbb{R} \) and \( n \in \mathbb{N} \), then there exists some \( p \in (0, 1) \) such that \( N \) is negative binomial with parameters \( (r, p) \), i.e.,

\[
f(n) = \binom{r}{n} p^r (1 - p)^n, \quad n \in \mathbb{N}.
\]

Moreover, for the same value of \( p \), \( N_j, j = 1, \ldots, k \), are independent negative binomials with parameters \( (r_j, p) \), \( j = 1, \ldots, k \), respectively (in the sense that \( N_j = 0 \) a.s., whenever \( r_j = 0 \)).

**Proof.** The assumptions of Theorem 2.1 are satisfied with

\[
a_j(n) = \binom{r_j}{n}, \quad j = 1, \ldots, k, \quad c(n) = \binom{r}{n} \quad \text{and} \quad b(n) = \binom{r - r_1 - r_2}{n}, \quad n \in \mathbb{N}.
\]

Moreover, if for some \( \theta > 0 \), \( \sum_n a_2(n) \theta^n < \infty \), then necessarily \( \theta < 1 \). Therefore, the linearly independent functions

\[
g_n(x) = \binom{x}{n}, \quad x \geq 0, \quad n \in \mathbb{N},
\]

satisfy assumptions (B1)–(B3) of Theorem 2.1 with \( M = 1 \) and \( a = r - r_1 - r_2 \geq 0 \). Thus, \( c < 1 \), \( p = 1 - c \in (0, 1) \) and the desired result follows. \( \square \)

**Corollary 2.3.** Suppose that the r.v.’s \( N_1, \ldots, N_k, k \geq 3 \), take values in \( \mathbb{N} \) and satisfy the RR condition (1.1), and assume that \( N = N_1 + \cdots + N_k \) has p.m.f. \( f(n) \), \( n \in \mathbb{N} \), with \( f(0) < 1 \). If there exist integers \( m_3 \geq 0, \ldots, m_k \geq 0 \) such that for every \( n \in \mathbb{N} \) with \( f(n) > 0 \),

\[
\mathbb{P}[N_1 = n_1, \ldots, N_k = n_k | N = n] = \left( \prod_{j=3}^{k} \binom{m_j}{n_j} \right) / \sum_{j=0}^{\min[m,n]} \binom{m}{j} (n + 1 - j)
\]

for all \( n_1, \ldots, n_k \in \mathbb{N} \) with \( n_1 + \cdots + n_k = n \), where \( m = m_3 + \cdots + m_k \geq 0 \) and

\[
\binom{x}{n} = \begin{cases}
1, & \text{if } n = 0, \\
\frac{1}{n!} \prod_{s=0}^{n-1} (x - s) & \text{if } n \geq 1
\end{cases}
\]

for all \( x \in \mathbb{R} \) and \( n \in \mathbb{N} \), then \( N_1, \ldots, N_k \) are independent and there exists some \( p \in (0, 1) \) such that \( N_1 \) and \( N_2 \) are Geometric \( (p) \) with \( \mathbb{P}[N_j = n_j] = p(1 - p)^{n_j}, n_j \in \mathbb{N}, j = 1, 2 \), while for \( j \geq 3 \), \( N_j \) is Binomial \( (m_j, p/(2 - p)) \) (in the sense that \( N_j = 0 \) a.s., whenever \( m_j = 0 \)).
Proof. The assumptions of Theorem 2.1 are satisfied with

\[
\begin{align*}
  a_1(n) &= a_2(n) = 1, \\
  a_j(n) &= \binom{m_j}{n}, \quad j = 3, \ldots, k, \\
  c(n) &= \min\{m, n\} \sum_{j=0}^\infty \binom{m}{j} (n + 1 - j) \\
  b(n) &= \binom{m}{n}, \quad n \in \mathbb{N}.
\end{align*}
\]

Moreover, if for some \( \theta > 0 \), \( \sum_n a_2(n) \theta^n < \infty \), then necessarily \( \theta < 1 \). Therefore, the linearly independent functions

\[
g_n(x) = \binom{x}{n}, \quad x \geq 0, \quad n \in \mathbb{N},
\]

satisfy assumptions (B1)–(B3) of Theorem 2.1 with \( M = 1 \) and \( a = m \geq 0 \). Thus, \( c < 1 \), \( p = 1 - c \in (0, 1) \) and the desired result follows. \( \square \)

References