On Characterizing Some Discrete Distributions Using an Extension of the Rao-Rubin Theorem

John Panaretos

1982

Online at http://mpra.ub.uni-muenchen.de/6229/
ON CHARACTERIZING SOME DISCRETE DISTRIBUTIONS
USING AN EXTENSION OF THE RAO-RUBIN
THEOREM

By JOHN PANARETOS*

Department of Statistics, Trinity College, Dublin, Ireland

1. INTRODUCTION

Let \( X, Y \) be non-negative integer-valued r.v.'s such that
\[
P(X = n) = P^n
\]
and
\[
P(Y = r | X = n) = s(r, n), \quad r = 0, 1, \ldots, n.
\]
Using Bernstein's theorem on completely monotone functions, Rao and Rubin (1964) have shown that if
\[
s(r, n) = \binom{n}{r} p^r q^{n-r}, \quad p \in (0, 1), q = 1-p
\]
then \( \{P_n, n = 0, 1, \ldots\} \) is Poisson if and only if (iff)
\[
P(Y = r) = P(Y = r | X = Y) = P(Y = r | X > Y), \quad r = 0, 1, \ldots
\]
(the so called R-R condition).

Shanbhag (1974) gave an elementary proof of this result. Srivastava and Srivastava (1970) proved that if \( \{P_n\} \) is Poisson, then the R-R condition holds iff \( s(r, n), \quad r = 0, 1, \ldots, n \) is binomial. Rao and Rubin (1964) have also shown that if the non-negative integer-valued r.v. \( X \) takes the values \( k, k+1, \ldots; k \geq 0 \) and \( s(r, n) \) is binomial as in (1.1), then \( \{P_n, n = k, k+1, \ldots\} \) is truncated Poisson iff
\[
P(Y = r | Y \geq k) = P(Y = r | X = Y) = P(Y = r | X > Y, Y \geq k),
\]
\[
r = k, k+1, \ldots
\]

*Present address: Department of Statistics of Actuarial Science, University of Iowa, Iowa City, Iowa, 52242, USA.
2. 

Shanbhag's result and a variant

Shanbhag (1977) gives the following extension of the R-R characterization.

Theorem 1 (Shanbhag's extension): Let \( \{ (a_n, b_n), n = 0, 1, \ldots \} \) be a sequence of real vectors with \( b_0, b_1 > 0 \), \( a_n > 0 \) for all \( n \geq 0 \) and \( b_n \geq 0 \) for \( n \geq 2 \) and let \( \{ c_n \} \) denote the convolution of \( \{ a_n \} \) and \( \{ b_n \} \). (Observe that \( c_n > 0 \), \( n \geq 0 \)). Let \( (X, Y) \) be a random vector, as defined in the introduction, such that \( P_0 = P[X = 0] < 1 \) and whenever \( P_n > 0 \)

\[
P(Y = r \mid X = n) = \frac{a_n b_n - r}{c_n}, \quad r = 0, 1, \ldots, n. \quad \ldots (2.1)
\]

Then

\[
P(Y = r) = P(Y = r \mid X = Y) \quad r = 0, 1, \ldots \quad \ldots (2.2)
\]

iff

\[
\frac{P_n}{c_n} = \frac{P_0}{c_0} \theta^n \quad n = 1, 2, \ldots \quad \text{for some } \theta > 0. \quad \ldots (2.3)
\]

The result of Theorem 1 will now be used to prove the following theorem.

Theorem 2 (Extension to the truncated case): Let \( \{ (a_n, b_n) : n = 0, 1, \ldots \} \) be a sequence of vectors of non-negative real numbers such that \( b_0, b_1 > 0 \) and \( a_n > 0 \) for \( n \geq k, \quad k \geq 0 \). Let \( \{ c_n \} \) be the convolution of \( \{ a_n \} \) and \( \{ b_n \} \). (Observe that \( c_n > 0, \quad n \geq k \)). Let \( (X, Y) \) be a vector of non-negative integer-valued r.v.'s such that \( P(X = n) = P_n \) with \( P_k < 1 \) and \( X \) taking values \( \geq k \) only, and whenever \( P_n > 0 \)

\[
P(Y = r \mid X = n) = \frac{a_n b_n - r}{c_n}, \quad r = 0, 1, \ldots, n
\]

\[
\quad n = k, k+1, \ldots \quad \ldots (2.4)
\]

Then

\[
P(Y = r \mid Y \geq k) = P(Y = r \mid X = Y), \quad r = k, k+1, \ldots \quad \ldots (2.5)
\]

iff

\[
\frac{P_n}{c_n} = \frac{P_k}{c_k} \theta^{n-k} \quad n = k, k+1, \ldots \quad \text{for some } \theta > 0. \quad \ldots (2.6)
\]

Proof: We have been given that \( X-k \) is a non-negative integer-valued random variable. Further, it follows that conditional on \( Y-k \geq 0 \), the random variable \( Y-k \) is non-negative and integer valued.
If we define

$$c_n^{(k)} = \sum_{r=k}^{n} a_r b_{n-r}$$  \hspace{1cm} \ldots \hspace{0.5cm} (2.7)$$

we will obviously have

$$c_{n+k}^{(k)} = \sum_{r=k}^{n+k} a_r b_{n+k-r} = \sum_{r=0}^{n} a_{r+k} b_{n-r}.$$  \hspace{1cm} \ldots \hspace{0.5cm} (2.8)$$

It then follows that

$$P(X-k = n \mid Y-k \geq 0) = \frac{1}{P(Y \geq k)} c_{n+k}^{(k)} \frac{P_{n+k}}{c_{n+k}} \hspace{0.5cm} n = 0, 1, \ldots \hspace{1cm} \ldots \hspace{0.5cm} (2.9)$$

and

$$P(Y-k = r \mid X-k = n, Y-k \geq 0) = \frac{a_{r+k} b_{n-r}}{c_{n+k}^{(k)} a_{n+k}} \hspace{0.5cm} r = 0, 1, \ldots, n \hspace{1cm} \ldots \hspace{0.5cm} (2.10)$$

It also follows that (2.5) is equivalent to

$$P(Y-k = r \mid Y-k \geq 0) = P(Y-k = r \mid X-k = Y-k) \hspace{0.5cm} r = 0, 1, \ldots \hspace{1cm} \ldots \hspace{0.5cm} (2.11)$$

Then, clearly the random vector $\langle X-k, Y-k \rangle$ conditional on $Y-k \geq 0$ possesses all the properties required of $(X, Y)$ in Theorem 1. Hence, applying this theorem to the random vector $\langle X-k, Y-k \rangle$ conditional on $Y-k \geq 0$ we have that (2.11) holds iff

$$P(X-k = n \mid Y-k \geq 0) = \frac{P(X-k = 0 \mid Y-k \geq 0) g_n}{c_{n+k}^{(k)}} \hspace{0.5cm} \text{for some } \theta > 0$$

which in view of (2.9), is equivalent to

$$\frac{P_{n+k}}{c_{n+k}} = \frac{P_{\theta} g_n}{c_{k}} \hspace{0.5cm} n = 0, 1, \ldots \hspace{1cm} \ldots \hspace{0.5cm} (2.12)$$

Since (2.11) is equivalent to (2.5), the result is now obvious.

**Note 1**: Theorem 2 can also be obtained directly by appealing to the basic lemma of Shaubhag (1977).

**Note 2**: It is interesting to observe that Theorems 1 and 2 remain valid if the right hand side or the left hand side of (2.2) and (2.5) is replaced by $P(Y = r \mid X > Y)$ and $P(Y = r \mid X > Y, Y \geq k)$ respectively.
In part 3, Theorem 1 is used to show that if $s(r, n)$ is negative hypergeometric $(n, m, \rho)$ then the R-R condition holds if $P_n$ is negative binomial $(m+\rho, p)$ without requiring additional conditions on the existence of the derivatives $G^{(n)}(t)$ of the p.g.f. $G$ of $X$, as in Patil and Ratnaparkhi (1975).

Then Theorem 2 is used to show that provided again that $s(r, n)$ is negative hypergeometric, the modified R-R condition (2.5) is necessary and sufficient for $P_n, n = k, k+1, \ldots$ to be truncated negative binomial.

In part 4 we state two other characterizations based on truncated forms of the distribution of $Y$ given $X$.

3. CHARACTERIZATIONS OF THE NEGATIVE BINOMIAL AND TRUNCATED NEGATIVE BINOMIAL DISTRIBUTIONS

Corollary 3.1 (Characterization of the Negative Binomial Distribution): Let $X$ and $Y$ be non-negative integer-valued r.v.'s as defined in Theorem 1.

Assume that

$$P(Y = r | X = n) = \binom{m}{r} \binom{-\rho}{n-r} \binom{-m-\rho}{n} r = 0, 1, \ldots, n \quad m > 0, \rho > 0 \quad \ldots \quad (3.1)$$

i.e., negative hypergeometric $(n, m, \rho)$.

Then the condition (2.2) holds iff

$$P_n = P(X = n) = \binom{-N}{n} p^N (-q)^n \quad \text{for } n = 0, 1, \ldots \quad \ldots \quad (3.2)$$

Proof: It is obvious that Theorem 1 applies with

$$a_n = \binom{-m}{n} (-1)^n, \quad b_n = \binom{-\rho}{n} (-1)^n, \quad \ldots \quad (3.3)$$

whence

$$c_n = \binom{-m-\rho}{n} (-1)^n \quad \ldots \quad (3.4)$$

Corollary 3.2 (Characterization of the truncated negative binomial distribution): Let us now assume that $X, Y$ are non-negative integer-valued r.v.'s as defined in Theorem 2. Suppose that the conditional distribution of
Y given X is negative hypergeometric as in (3.1). Then, the modified R-R condition (2.5) holds iff

\[ P_n = \frac{\binom{-N}{n} p^n q^n}{\sum_{i=0}^{n} \binom{-N}{i} p^i q^i} \quad n = k, k+1, \ldots \quad \ldots \quad (3.5) \]

i.e., negative binomial truncated at \( k-1 \).

**Proof**: Define the sequences \( \{a_n\}, \{b_n\}, n = 0, 1, \ldots \) as in (3.3). Consequently \( P(Y = r \mid X = n) \) is again of the form \( a_r b_{n-r}/c_n \), \( r = 0, 1, \ldots, n \).

Hence, since all the requirements of Theorem 2, are met, (2.5) holds iff \( P_n = P_k \sum c_k \theta^k \) for some \( \theta > 0 \), and \( n = k, k+1, \ldots \) i.e., iff (3.5) is valid.

**Note**: It is interesting to note here that if \( \{P_n\} \) is negative binomial and \( P(Y = r \mid X = n) \) is of the form \( a_r b_{n-r}/c_n \), then the negative hypergeometric is not the only distribution of the form (2.1) that \( Y \mid X \) should follow in order that the R-R condition (2.2) holds.

This is so, because there exist two independent non-negative random variables which are not negative binomial, but their sum is negative binomial. However it is interesting to observe that if we also require \( \{b_n\} \) to be the s-fold convolution of \( \{a_n\} \), then we get a characterization for the negative hypergeometric distribution.

4. **Characterizations of the convolution of distributions with one of the members as truncated**

Corollary 4.1 (Characterization of the convolution of a Poisson \((\mu)\), with a truncated Poisson \((\lambda)\) Distribution): Consider the r.v.'s \( X \) and \( Y \) as in Theorem 2. Suppose that the conditional distribution of \( Y \mid X \) is binomial truncated at \( k-1 \), i.e.,

\[ P(Y = r \mid X = n) = \binom{n}{r} p^r q^{n-r} \quad r = k, k+1, \ldots, n \]

\[ \sum_{i=k}^{n} \binom{n}{i} p^i q^{n-i} \quad 0 < p < 1, q = 1-p \quad \ldots \quad (4.1) \]

Then, the modified R-R condition (2.5) holds iff

\[ P(X = n) = \frac{e^{-\mu} \sum_{r=k}^{\infty} \binom{n}{r} \lambda^r \mu^{n-r}}{n! \sum_{i=k}^{\infty} \lambda^i \mu^{n-i}} \quad \text{for some } \lambda, \mu > 0 \quad \ldots \quad (4.2) \]

i.e., convolution of a Poisson \((\mu)\) with a Poisson \((\lambda)\) truncated at \( k-1 \).
Proof: Consider the following sequences,

\[ a_n = \begin{cases} 
\frac{\lambda^n/n!}{\sum_{i=1}^{n-1} \lambda^i/i!} & n = k, k+1, \ldots \\
0 & n = 0, 1, \ldots, k-1 
\end{cases} \quad \ldots \quad (4.3) \]

\[ b_n = \frac{e^{-\mu} \mu^n}{n!} \quad n = 0, 1, \ldots \quad \ldots \quad (4.4) \]

The convolution of (4.3) and (4.4) for \( n \geq k \) is given by

\[ c_n = \sum_{r=k}^{n} a_r b_{n-r} = \frac{e^{-\mu} \sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r}}{n! \sum_{i=2}^{n} \lambda^i/i!} \quad \ldots \quad (4.5) \]

It is clear now that the R.H.S. of (4.1) can be written as \( a_n b_{n-k}/c_n \) with \( a_n, b_n, c_n \) given by (4.3), (4.4), (4.5) respectively and with \( p = \frac{\lambda}{\lambda+\mu}, \quad q = 1-p \).

Since \( a_n, b_n \) satisfy all the conditions of Theorem 2 we have that the R.R. condition (2.5) holds iff \( P_n \) satisfies (2.6) for \( c_n \) given by (4.5). This implies that (2.5) holds iff

\[ P_n = P_k \frac{\sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r}/n!}{\binom{k}{k} \lambda^k/k!} \quad \text{if } n = k, k+1, \ldots \text{ for some } \theta > 0. \]

Noting that \( \sum_{n=k}^{\infty} P_n = 1 \) we have

\[ P_k = e^{-\mu \theta} \frac{\sum_{r=k}^{\infty} (\lambda \theta)^r/r!}{(\lambda \theta)^k/k!} \]

So, (2.5) holds iff

\[ P_n = e^{-\mu \theta} \frac{\sum_{r=k}^{n} (\lambda \theta)^r (\mu \theta)^{n-r}}{n! \sum_{r=k}^{\infty} (\lambda \theta)^r/r!} \quad \lambda, \mu, \theta > 0 \]

i.e. iff \( \{P_n\} \) is the probability distribution of the convolution of a Poisson r.v. with a \( k-1 \) truncated Poisson r.v.
Note: The form of the convoluted Poisson distribution examined in Corollary 4.1 is a special case of convoluted Poisson distributions studied by Samaniego (1976).

Corollary 4.2 (Characterization of the convolution of a negative binomial with a truncated negative binomial distribution): Let again, $X, Y$ be as in Theorem 2. Suppose that the distribution of $Y | X$ is negative hypergeometric truncated at $k-1$, i.e.,

$$P(Y = r | X = n) = \binom{-m}{r} \binom{-\rho}{n-r} \frac{\sum_{i=k}^{n} \binom{-m}{i} \binom{-\rho}{n-i}}{\sum_{i=k}^{n} \binom{-m}{i} \binom{-\rho}{n-i}} r = k, k+1, \ldots, n \ldots \quad (4.6)$$

In that case, the modified R-R condition (2.5) holds iff $\{P_n\}$ is the convolution of a negative binomial $(\rho, p)$ and a negative binomial $(m, p)$ truncated at the point $k-1$, i.e., iff

$$P_n = \sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-q)^{n-r} p^r \frac{\sum_{i=k}^{n} \binom{-m}{i} \binom{-\rho}{n-i}}{\sum_{i=k}^{n} \binom{-m}{i} \binom{-\rho}{n-i}} n = k, k+1, \ldots, n \ldots \quad (4.7)$$

Proof: Define the sequences

$$a_n = \begin{cases} \sum_{i=k}^{n} \binom{m+n-1}{r} \binom{n}{r-i} q^{n-i} & n = k, k+1, \ldots, n \ldots \quad (4.8) \\ 0 & n = 0, 1, \ldots, k-1 \end{cases}$$

$$b_n = \binom{\rho+n-1}{n} q^n \quad n = 0, 1, \ldots \quad (4.9)$$

Then, the convolution of $\{a_n\}$ and $\{b_n\}$ is given by

$$c_n = \sum_{r=k}^{n} \binom{m+r-1}{r} \binom{\rho+n-r-1}{n-r} q^n \frac{\sum_{i=k}^{n} \binom{m+r-1}{i} \binom{\rho+n-r-1}{n-r}}{\sum_{i=k}^{n} \binom{m+r-1}{i} \binom{\rho+n-r-1}{n-r}} n = k, k+1, \ldots \ldots \quad (4.10)$$

Clearly, since $\binom{m+n-1}{n} = \binom{-m}{n} (-1)^n$

$$\frac{a_n b_{n-r}}{c_n} = \frac{\binom{-m}{r} \binom{-\rho}{n-r}}{\sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r}} r = k, k+1, \ldots \ldots \quad (4.11)$$
Consequently, the sequences \( \{a_n\}, \{b_n\} \) as defined in (4.8), (4.9) can be used to express (4.6) in the form required by Theorem 2. Hence the result follows readily from Theorem 2.

**Note:** It is clear that for \( k = 0 \) Corollaries 4.1 and 4.2 reduce to Corollaries 3.1 and 3.2, respectively.

**References**


*Paper received: March, 1979.*

*Revised: April, 1980.*