CHAPTER 3.

SHANBHAG'S EXTENSION OF THE R-R CHARACTERIZATION

THE UNIVARIATE CASE

3.0 Introduction

In Chapter 2, we examined various characterizations, based on the R-R condition. However, the techniques used require lengthy proofs.

Shanbhag (1976) introduces an Extension of the R-R characterization; this characterizes a class of statistical distributions using the R-R condition. It also provides a simple way of characterizing a number of distributions as survival distributions or as original distributions. In this chapter, we state Shanbhag's Extension, and we utilize it to arrive at some characterizations. We also use it to provide simpler proofs for some of the results given in Chapter 2 and also for variants of some others. An extension of his result is derived enabling us to get characterizations of the truncated forms of the distributions which were characterized previously.

Finally, a somewhat different form of the extension is introduced, and characterization based on truncated forms of survival distributions are derived.

3.1 Shanbhag's Extension

Lemma 3.1.1 (Shanbhag 1976).

Let $\{(V_n, W_n)\ n=0,1,\ldots\}$ be a sequence of vectors with non-negative real components such that

$$V_n \neq 0 \text{ for some } n \geq 1$$

and

$$W_1 \neq 0.$$

Then,

$$V_m = \sum_{n=0}^{\infty} V_{n+m} W_n, \qquad m=0,1,\cdots, \tag{3.1.1}$$

iff

$$\sum_{n=0}^{\infty} W_n b^n = 1 \qquad (3.1.2a)$$

and

$$V_n = V_0 b^n \quad n=1,2,\ldots \qquad (3.1.2b)$$

for some $b > 0$.

The proof of the lemma is given by Shanbhag; it is based on a technique used in renewal theory. Assuming that (3.1.1) holds he proves that $V_m = 0$ iff $V_n = 0$ for all $n > m$. Then using that, he establishes that $V_n \neq 0$ for every $n \geq 0$. Defining $b = \sup \{V_m^* \; m \geq 1\}$ where $V_m^* = \frac{V_m}{V_{m-1}}$ $m=1,2,\ldots$ he shows that b is in fact equal to V_m^*, and hence (3.1.2b) is valid.

Applying lemma (3.1.1) Shanbhag arrives at the following.

<u>Theorem 3.1.1</u> (Shanbhag's extension of the R-R characterization)

Let $\{(a_n, b_n): n=0,1,\ldots\}$ be a sequence of real vectors with

$$a_n > 0 \text{ for every } n \geq 0$$

and

$$b_0 > 0,\; b_1 > 0,\; b_n \geq 0 \text{ for } n \geq 2\;.$$

Denote by $\{c_n\}$ the convolution of $\{a_n\}$ and $\{b_n\}$. Let (X,Y) be a random vector of non-negative integer-valued components such that

$$P\{X=n\} = P_n \quad n \geq 0 \text{ with } P_0 < 1 \qquad (3.1.5)$$

and whenever $P_n > 0$ we have

$$P(Y=r|X=n) = \frac{a_r\, b_{n-r}}{c_n}, \quad r=0,1,\ldots,n. \qquad (3.1.6)$$

Then

$$P(Y=r) = P(Y=r|X=Y) \quad r=0,1,\ldots \qquad (3.1.7)$$

iff

$$\frac{P_n}{c_n} = \frac{P_0}{c_0}\,\theta^n \quad n=1,2,\ldots \quad \text{for some } \theta > 0. \tag{3.1.8}$$

<u>Proof</u>

Consider the sequences V_n, W_n as follows

$$V_n = \frac{P_n}{c_n} \quad n \geq 0 \tag{3.1.9}$$

$$W_n = b_n \left\{ \sum_{n=0}^{\infty} P_n \frac{a_n}{c_n} \right\}. \tag{3.1.10}$$

Then the above defined sequences satisfy the requirements of lemma 3.1.1.

Obviously V_n, W_n have non-negative real components $V_n \neq 0$ for some $n \geq 1$ ($c_n > 0$ and if P_n were zero for all $n \geq 1$ then P_0 would be 1, a contradiction). Also, since $b_1 > 0$ and $\sum_{n=0}^{\infty} P_n \frac{a_n}{c_n} \neq 0$ we have that $W_1 \neq 0$

It is clear that

$$V_n = V_0 b^n \text{ is equivalent to } \frac{P_n}{c_n} = \frac{P_0}{c_0}\,\theta^n$$

On the other hand, it can be checked that $P(Y=r) = P(Y=r|X=Y)$ is equivalent to the formula (3.1.1) of the lemma 3.1.1. So, (3.1.7) is valid iff (3.1.8) is true.

<u>Note 1</u>. It is worth noting here that for the above theorem, the relation (3.1.2a) is redundant.

<u>Note 2</u>. The above theorem remains valid if the R.H.S. or the L.H.S. of (3.1.7) is replaced by $P(Y=r|X>Y)$.

<u>Note 3</u>. It is interesting to observe that if the vector (X,Y) is such that (3.1.8) is valid, then the r.v.'s Y and $Y-X$ are independent.

This is so, because,

$$P(Y=r|X=n) = \frac{P(Y=r,X=n)}{P(X=n)} = \frac{P(Y=r,X-Y=n-r)}{P_n}$$

and using (3.1.6) we have

$$\frac{P(Y=r, X-Y=n-r)}{P_n} = \frac{a_r b_{n-r}}{c_n} \quad \text{or } \left(\text{from } (3.1.8)\right)$$

$$P(Y=r, X-Y=n-r) = \frac{P_0}{c_0} a_r \theta^r b_{n-r} \theta^{n-r} \quad \text{i.e.,}$$

$$P(Y=r, X-Y=n) = \frac{P_0}{c_0} a_r \theta^r b_n \theta^n \quad . \qquad (3.1.11)$$

Hence the result follows.

3.2 Shanbhag's Extension in Relation to a Theorem by Patil and Seshadri

Patil and Seshadri (1964) (also in K.L.R. p.424) prove the following theorem.

Theorem 3.2.1 (Patil and Seshadri (1964).) Let Y and Z=X-Y be independent discrete r.v's and let $P(Y=r|X=n) = s(r,n)$.

Then, if $s(r,n)$ satisfies the relation

$$\frac{s(n,n)s(0,n-r)}{s(r,n)s(n-r,n-r)} = \frac{h(n)}{h(r)h(n-r)} \qquad (3.2.1)$$

where $h(x)$ is a non-negative function, then the distributions of Y and $Z = X-Y$ are given by

$$P(Y=r) = P(Y=0)h(r)e^{ar} \qquad (3.2.2)$$

$$P(Z=z) = P(Z=0)k(z)e^{az} \qquad (3.2.3)$$

with

$$k(z) = \frac{h(z)s(0,z)}{s(z,z)} \; . \qquad (3.2.4)$$

Patil and Ratnaparkhi(1975) observe that

$$P(X=n) = P(X=0)f(n)e^{an} \qquad (3.2.5)$$

where

$$f(n) = \frac{h(n)}{s(n,n)} \; . \qquad (3.2.6)$$

This latter result follows readily from the above mentioned theorem of Patil and Seshadri.

Looking at this theorem as far as the distribution of X is concerned, one can see that in terms of the terminology we have used in Shanbhag's Extension, the result confirms that if Y and X-Y are independent and s(r,n) is of a form similar to (3.1.6) then,

$$\frac{P_n}{c_n} = \frac{P_0}{c_0}\theta^n \quad n = 1,2,\ldots \text{ for some } \theta>0. \tag{3.2.7}$$

However, Patil and Seshadri's theorem is more restricted than the "only if" part of Shanbhag's Extension since it requires independence of Y and X-Y, whereas Shanbhag's assumes that Y and X-Y are independent only over the set X-Y = 0. In addition, by taking into consideration Note 3 of section 3.1 we can see that relation (3.1.8) is necessary and sufficient for Y and X-Y to be independent. Patil and Seshadri's theorem provides only the "sufficient" part of that result.

3.3 Some Characterizations Based on the Extension

As it is pointed out by Shanbhag (1976), theorem 2.1.1 (R-R characterization of the Poisson distribution) and a variant of theorem 2.3.1 (Srivastava and Srivastava) can be obtained in a simpler way.

In fact, if we define

$$a_n = \frac{p^n}{n!}, \quad b_n = \frac{(1-p)^n}{n!} \quad n=0,1,\ldots \tag{3.3.1}$$

then the binomial distribution can be written in the form (3.1.6) (since for a_n and b_n as defined in (3.3.1), $c_n = \frac{1}{n!}$ n=0,1,...). This means that according to the theorem (3.1.1) the R-R condition (3.1.7) holds iff

$$P_n = P_0 \frac{\lambda^n}{n!} \quad \lambda > 0,$$

which gives P_n as Poisson (λ).

The following is the variant of Srivastava and Srivastava's theorem 2.3.1.

Corollary 3.3.1 (Shanbhag 1976)

If (X,Y) is a random vector of non-negative integer-valued components such that (3.1.6) is satisfied and P_n is Poisson (λ) then (3.1.7) holds iff

$$P(Y=r|X=n) = \binom{n}{r} p^r (1-p)^{n-r}$$

$$0 \leqslant r \leqslant n, \quad n \geqslant 0 \qquad (3.3.2)$$

for some $p \in (0,1)$.

Proof

Since the requirements of theorem (3.1.1) are met we have that (3.1.7) holds iff

$$\frac{P_n}{c_n} = \frac{P_0}{c_0}\theta^n, \text{ i.e. iff } c_n = \frac{P_n}{P_0}\theta^{-n}$$

which implies that

$$c_n = c_0 \frac{\mu^n}{n!} \qquad \left(\mu = \frac{\lambda}{\theta}\right).$$

But

$$c_n = \sum_{r=0}^{n} a_r b_{n-r} .$$

So

$$\sum_{r=0}^{n} a_r b_{n-r} = c_0 \frac{\mu^n}{n!}, \qquad \text{i.e. } \sum_{n=0}^{\infty}\sum_{r=0}^{n} a_r b_{n-r} = c_0 \sum_{n=0}^{\infty} \frac{\mu^n}{n!}$$

which gives

$$c_0 = \sum_{n=0}^{\infty}\sum_{r=0}^{n} a_r b_{n-r}\, e^{-\mu}$$

i.e.,

$$c_n = e^{-\mu}\frac{\mu^n}{n!}\sum_{m=0}^{\infty}\sum_{k=0}^{m} a_k b_{m-k},$$

$$\sum_{r=0}^{n} a_r b_{n-r} = e^{-\mu}\frac{\mu^n}{n!}\sum_{m=0}^{\infty}\sum_{k=0}^{m} a_k b_{m-k}.$$

Hence

$$\sum_{r=0}^{n} \frac{a_r}{\sum_{k=0}^{\infty} a_k} \frac{b_{n-r}}{\sum_{m=k}^{\infty} b_{m-k}} = e^{-\mu} \frac{\mu^n}{n!}$$

and by making use of Raikov's theorem (1937) we come to the conclusion that (3.1.7) holds iff

$$\frac{a_n}{\sum_{m=0}^{\infty} a_m} = e^{-\mu_1} \frac{\mu_1^n}{n!} \qquad (3.3.3)$$

$$\frac{b_n}{\sum_{m=0}^{\infty} b_m} = e^{-\mu_2} \frac{\mu_2^n}{n!} \qquad (3.3.4)$$

with $\mu_1+\mu_2 = \mu$.

Furthermore, since we have assumed that $P(Y=r|X=n)$ is such that (3.1.6) is satisfied, then taking a_n and b_n from (3.3.3) and (3.3.4) we have that

$$P(Y=r|X=n) = \binom{n}{r} p^r q^{n-r} \qquad (3.3.5)$$

with $p = \frac{\lambda_1}{\lambda}$ and $p \in (0,1)$.

This gives the "only if" part of the corollary. The "if" part has been considered in our earlier remarks.

Corollary 3.3.2 (Characterization of the Negative Binomial)

Let (X,Y) be a random vector on non-negative integer-valued components such that

$$P(Y=r|X=n) = \frac{\binom{-m}{r}\binom{-\rho}{n-r}}{\binom{-m-\rho}{n}} \qquad r = 0,1,\ldots,n \quad m>0,\ \rho>0 \qquad (3.3.6)$$

i.e. Neg. Hypergeometric (n,m,ρ).

Then the R-R condition (3.1.7) holds iff

$$P_n = \binom{-N}{n} p^N (-q)^n \qquad N=m+\rho \tag{3.3.7}$$

i.e. Negative Binomial with parameters ρ, $m+\rho$.

Proof

Define

$$a_n = \binom{m+n-1}{n} q^n \quad b_n = \binom{\rho+n-1}{n} q^n \ . \tag{3.3.8}$$

Then,

$$C_n = \sum_{r=0}^{n} a_r b_{n-r} = \sum_{r=0}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-q)^n = \binom{m+\rho+n-1}{n} q^n$$

Then, obviously, (3.3.6) is of the form (3.1.6) with a_n, b_n as in (3.3.8).

Hence, according to theorem (3.1.1), the R-R condition holds iff

$$P_n = P_0 \frac{c_n}{c_0} \theta^n \text{ for some } \theta > 0$$

which finally gives

$$P_n = \binom{-N}{n} p^N (-q)^n \qquad N = \rho+m.$$

Note 1

Corollary (3.3.2) is an improved version of a result appearing in Patil and Ratnaparkhi(1975).

They make the same characterization using an extension of Patil and Seshadri's theorem, but under the additional assumption that the r-th order derivative $G_X^{(r)}(t)$ r=1,2,.. of the p.g.f. of X, exists at the point 1. This assumption is not required in our proof.

Note 2

It must be noted here, that if P_n is Negative Binomial and $P(Y=r|X=n)$ is of the form $a_r b_{n-r}/c_n$, then the Negative Hypergeometric is not the only distribution that $Y|X$ should follow in order that the R-R condition holds.

This is so, because there exist two independent non-negative random variables which are not Negative Binomial, but whose sum is Negative Binomial. Specifically the p.g.f. of the Negative Binomial distribution can be written as

$$G(t) = \left(\frac{p}{1-qt}\right)^N = \exp N\{\log(1-q)-\log(1-qt)\}$$

$$= \exp N\left\{qt + \frac{q^2t^2}{2} + \ldots - q - \frac{q^2}{2} - \frac{q^3}{3} - \ldots\right\}$$

$$= \exp N \sum_{i=1}^{\infty} \frac{q^i(t^i-1)}{i} .$$

But

$$\exp\left\{\frac{Nq^i}{i}(t^i-1)\right\}$$

is the p.g.f. of the r.v. iX, $X \sim$ Poisson $\left(N\frac{q^i}{i}\right)$. This implies that the Negative Binomial may be obtained as the convolution of two random variables, one of which is i times Poisson with parameter $N\frac{q^i}{i}$.

However, it is interesting to observe that if we also require $\{b_n\}$ to be the s-fold convolution of $\{a_n\}$, then the difficulty mentioned previously is overcome and we get a characterization for the Negative Hypergeometric as follows.

Corollary 3.3.3 (Characterization of the Negative Hypergeometric.)

Let (X,Y) be a random vector as in Corollary 3.3.2. Suppose that the conditional distribution of $Y|X$ is of the form $a_r b_{n-r}/c_n$ with the sequence $\{b_n\}$ being the s-fold convolution of $\{a_n\}$ (s fixed). Suppose also that the distribution of X is Negative Binomial. Then the R-R condition (3.1.7) holds iff $P(Y=r|X=n)$ is Negative Hypergeometric as in (3.3.6).

Proof This follows by using the same steps as in Corollary 3.3.1.

Remark

Corollary 3.3.1 is a variant of theorem 2.3.1. The difference being, that while in theorem 2.3.1 the result is restricted to the case where the parameter λ of the Poisson distribution is a variable, in 3.3.1 is valid for λ fixed.

Of course that was made possible by imposing the additional condition that $P(Y=r|X=n)$ is of the form $a_r b_{n-r}/c_n$.

3.4 An Interesting Limiting Case.

Let us consider the situation where the conditional distribution of $Y|X$ is of the form

$$P^{(\varepsilon)}(Y=r|X=n) = \frac{\left\{(1-\varepsilon)\binom{m}{r} + \varepsilon \frac{p^r}{r!}\right\}\left\{(1-\varepsilon)\binom{N-m}{n-r} + \varepsilon \frac{(1-p)^{n-r}}{(n-r)!}\right\}}{\sum_{j=0}^{n}\left\{(1-\varepsilon)\binom{m}{j} + \varepsilon \frac{p^j}{j!}\right\}\left\{(1-\varepsilon)\binom{N-m}{n-j} + \varepsilon \frac{(1-p)^{n-j}}{(n-j)!}\right\}} \tag{3.4.1}$$

$0 < \varepsilon < 1.$

Then it can be shown that the R-R condition (3.1.7) holds iff

$$P_n^{(\varepsilon)} = \frac{(1-\varepsilon)^2\binom{N}{n} + \varepsilon^2 \frac{1}{n!} + \varepsilon(1-\varepsilon)\sum_{r=0}^{n}\left\{\binom{m}{r}\frac{(1-p)^{n-r}}{(n-r)!} + \binom{N-m}{n-r}\frac{p^r}{r!}\right\}\theta^n}{(1-\varepsilon)^2(1+\theta)^N + \varepsilon\, e^{\theta} + \sum_{n=0}^{\infty}\sum_{r=0}^{n}\left\{\binom{m}{r}\frac{(1-p)^{n-r}}{(n-r)!} + \binom{N-m}{n-r}\frac{p^r}{r!}\right\}} \tag{3.4.2}$$

for some $\theta > 0$.

Proof Consider the sequences

$$a_r^{(\varepsilon)} = (1-\varepsilon)\binom{m}{r} + \varepsilon\frac{p^r}{r!} \tag{3.4.3}$$

$$b_n^{(\varepsilon)} = (1-\varepsilon)\binom{N-m}{n} + \varepsilon\frac{(1-p)^n}{n!} \tag{3.4.4}$$

Then the convolution of $a_n^{(\varepsilon)}$ and $b_n^{(\varepsilon)}$ is

$$c_n^{(\varepsilon)} = (1-\varepsilon)^2\binom{N}{n} + \varepsilon^2\frac{1}{n!} + \varepsilon(1-\varepsilon)\sum_{r=0}^{n}\left\{\binom{m}{r}\frac{(1-p)^{n-r}}{(n-r)!} + \binom{N-m}{n-r}\frac{p^r}{r!}\right\} \tag{3.4.5}$$

Since the sequences $a_n^{(\varepsilon)}$, $b_n^{(\varepsilon)}$ can be used to express (3.4.1) in the form of (3.1.6), from theorem (3.1.1) we have that the R-R condition is valid iff (3.1.8) is satisfied, which gives as solution a probability function of the form (3.4.2) for some $\theta > 0$.

For $\varepsilon \simeq 1$ (3.4.1) gets close to the Binomial distribution and (3.4.2) to the Poisson. This conclusion supports the findings in the limiting case.

In the situation however where $\varepsilon \simeq 0$, and so, (3.4.1) gets close to the hypergeometric and (3.4.2) to the binomial, the result is not valid any longer.

This is so, because for $\varepsilon \simeq 0$ the sequence $a_r^{(\varepsilon)}$ given in (3.4.3) becomes $a_r = \binom{m}{r}$ which obviously is not positive for every r as it is required in theorem (3.1.1).

In other words, if the distribution of $Y|X$ is hypergeometric, the R-R condition is not sufficient for X to be binomial. (Necessity has been proved by Patil and Ratnaparkhi (1975)). This answers in the negative a relevant question which was put by Patil and Ratnaparkhi. (A counter-example showing this is given on page 104.)

In Chapter 5 it will be proved that there exists a variant of the R-R condition which is necessary and sufficient for X to be Binomial, when $Y|X$ is Hypergeometric.

The conclusion here is that a result that holds for a particular situation, is not valid any more when we consider its limiting case.

3.5 The Truncated Case of the Extension

Theorem 3.5.1

Let (X,Y) be a random vector of non-negative, integer-valued components and let k be a non-negative integer such that

$$P(X \geqslant k) = 1 \quad \text{and} \quad P(X > k) > 0.$$

Then, provided that we can find a sequence of real vectors $\{(a_n, b_n)\ n=0,1,2,\ldots\}$ such that

$$\begin{aligned} &a_n > 0 \text{ for every } n \geqslant 0 \\ &b_0 > 0,\ b_1 > 0,\ b_n \geqslant 0 \quad n \geqslant 2 \end{aligned} \qquad (3.5.1)$$

and

$$P(Y=r|X=n) = \frac{a_r b_{n-r}}{c_n} \quad r=0,1,\ldots,n \qquad (3.5.2)$$

whenever $P_n > 0$, where $\{c_n\}$ denotes the convolution of $\{a_n\}$ and $\{b_n\}$, we will have

$$P(Y=r|Y \geqslant k) = P(Y=r|X=Y) \quad r=k,k+1,\ldots \qquad (3.5.3)$$

iff

$$\frac{P_n}{c_n} = \frac{P_k}{c_k}\,\theta^{n-k} \quad \text{for some } \theta > 0 \quad n=k,k+1,\ldots \qquad (3.5.4)$$

Proof Define the sequence V_n and W_n as follows

$$V_{n-k} = \frac{P_n}{c_n} \quad n=k,k+1,\ldots \qquad (3.5.5)$$

$$W_n = \frac{\sum_{i=k}^{\infty} P_i \frac{a_i}{c_i}}{P(Y\geq k)} b_n \qquad n=0,1,\ldots \tag{3.5.6}$$

Obviously $V_m \neq 0$ for some $m \geq 1$ and $W_1 \neq 0$. We find that

$$\sum_{n=0}^{\infty} V_{n+m}W_n = V_m \quad \text{is equivalent to} \quad \sum_{n=0}^{\infty} V_{n+m-k}W_n = V_{m-k}$$

which in turn, by making use of (3.5.5) and (3.5.6) is equivalent to

$$\sum_{n=0}^{\infty} \frac{P_{n+m}}{c_{n+m}} \sum_{i=k}^{\infty} \frac{P_i \frac{a_i}{c_i} b_m}{P(Y\geq k)} = \frac{P_m}{c_m} \qquad m=k,k+1,\ldots \tag{3.5.7}$$

On the other hand,

$$\frac{P(Y=r)}{P(Y\geq k)} = P(Y=r|X=Y)$$

is (by means of 3.5.2) equivalent to

$$\frac{\sum_{n=r}^{\infty} P_n \frac{a_r b_{n-r}}{c_n}}{P(Y\geq k)} = \frac{P_r \frac{a_r b_0}{c_r}}{\sum_{i=k}^{\infty} P_i \frac{a_i b_0}{c_i}} \qquad \text{i.e. to}$$

$$\frac{\sum_{n=0}^{\infty} \frac{P_{n+r}}{c_{n+r}} b_n \sum_{i=k}^{\infty} P_i \frac{a_i}{c_i}}{P(Y\geq k)} = \frac{P_r}{c_r} \qquad r=k,k+1,\ldots \tag{3.5.8}$$

Since (3.5.7) and (3.5.8) are identical,

$$\sum_{n=0}^{\infty} V_{n+m}W_n = V_m \quad \text{is equivalent to} \quad \frac{P(Y=r)}{P(Y\geq k)} = P(Y=r|X=Y)$$

and so, taking into consideration Lemma 3.1.1,we come to the conclusion that (3.5.3) holds iff (3.5.4) holds.

Here again, it can be seen that if (3.5.2), (3.5.4) are valid then the r.v's Y truncated at k-1 and X-Y are independent. As a result of that, it follows that the relation $\sum_{n=0}^{\infty} W_n b^n = 1$ of the lemma is again redundant.

Corollary 3.5.1

Using the result of the previous theorem, a simpler way of proving the characterization of the truncated Poisson distribution given in 2.2 can be derived as a special case. This can be achieved by considering the sequences $\{a_n\}$ and $\{b_n\}$ as in (3.3.1).

Since the binomial distribution can be expressed in the form (3.2.2) it follows from theorem 3.5.1 that (3.5.3) is true iff

$$P_n = P_k \frac{c_n}{c_k} \theta^{n-k} \qquad n=k,k+1,\ldots$$

which in our case gives,

$$P_n = P_k \frac{k!}{n!} \theta^{n-k}$$

or finally

$$P_n = \frac{\frac{\lambda^n}{n!}}{\sum_{n=k}^{\infty} \frac{\lambda^n}{n!}} \qquad n=k,k+1,\ldots$$

Corollary 3.5.2 (Characterization of the Truncated Negative Binomial.)

Define X, Y and k as in theorem 3.5.1.

Suppose that the conditional distribution of $Y|X$ is Negative Hypergeometric as it is given in (3.3.3). Then the R-R condition for the truncated case, namely (3.5.3), holds iff

$$P_n = \frac{\binom{-N}{n} p^N (-q)^n}{\sum_{j=k}^{\infty} \binom{-N}{j} p^N (-q)^j} \qquad \begin{matrix} n=k,k+1,\ldots \\ N=m+\rho \end{matrix} \qquad (3.5.9)$$

These are the probabilities for the truncated negative Binomial distribution.

Proof If we define the sequence $\{a_n\}$, $\{b_n\}$ n=0,1,... as in (3.3.8), then,as explained already,(3.5.3) will be satisfied iff (3.5.4) is satisfied, which in fact gives (3.5.9) as solution.

3.6 A Remark on Shanbhag's Extension

As seen earlier, Shanbhag's extension of the R-R theorem, requires the existence of a sequence $\{(a_n, b_n): n=0,1,\ldots\}$ of real vectors with $a_n > 0$ for every $n \geq 0$, $b_0 > 0$, $b_1 > 0$ with $b_n \geq 0$ for $n \geq 2$, such that the conditional distribution of Y on X is of the form (3.1.6). The same conditions are required in the truncated case.

It can however be seen that the truncated version of the extension remains valid if the sequence $\{a_n\}$ is defined for $n \geq k$ where k is a positive integer, in such a way that $a_n > 0$ for all $n \geq k$, and the conditions on b_n remain the same. This is so, because even for a_n defined only for n=k,k+1,... the conditions set up by lemma 3.1.1 are still met; V_{m-k} and W_n continue to be defined as in (3.5.5) and (3.5.6) respectively. In the latter case we define

$$c_n = \sum_{r=k}^{n} a_r b_{n-r} \qquad n=k,k+1,\ldots.$$

In fact we have again

$$V_{m-k} = \frac{P_m}{c_m} > 0 \qquad m=k,k+1\ldots,$$

since $c_n > 0$ for all $n \geq k$ $\left(c_n = a_k b_0 + \sum_{r=k+1}^{n} a_r b_{n-r} > 0\right)$, and $P(X > k) > 0$ as required in theorem 3.5.1.

On the other hand

$$W_1 = \frac{\sum_{i=k}^{\infty} P_i \frac{a_i}{c_i}}{P(Y \geq k)} \; b_1 > 0$$

(since $P_i > 0$, for some i, $P(Y \geq k) > 0$, $b_1 > 0$). Taking this into consideration, the following version of theorem (3.5.1) can be established.

Theorem 3.6.1

Let X and Y be as in theorem (3.5.1). Let also $\{a_n\}$ n=k,k+1,... be a sequence of real numbers such that

$$a_n > 0 \text{ for all } n \geq k$$

and $\{b_n\}$ n=0,1,... be as defined by (3.5.1). Then, if whenever $P_n > 0$ n=k,k+1...

$$P(Y=r|X=n) = \frac{a_r b_{n-r}}{c_n} \qquad r=k,k+1,..., \tag{3.6.1}$$

we will find that (3.5.3) holds iff (3.5.4) holds.

The following corollaries can now be proved.

Corollary 3.6.1 (Characterization of the Distribution which is the Convolution of a Poisson with a Truncated Poisson.)

Suppose that the conditional distribution of Y on X (where Y, X are as in theorem 3.6.1) is truncated binomial, i.e. that

$$P(Y=r|X=n) = \frac{\binom{n}{r} p^r q^{n-r}}{\sum_{r=k}^{n} \binom{n}{r} p^r q^{n-r}} \qquad r=k,k+1,.... \tag{3.6.2}$$

Then the R-R condition (3.5.3) holds iff P_n arises from the convolution of a Poisson distribution (with parameter μ) with a truncated Poisson distribution (with parameter λ and truncation at the point k-1), in other words iff

$$P_n = \frac{e^{-\mu} \sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r}}{n! \sum_{n=k}^{\infty} \frac{\lambda^n}{n!}} \qquad n=k,k+1,... . \tag{3.6.3}$$

Proof Let

$$a_r = \frac{\lambda^r}{r!} \Bigg/ \sum_{n=k}^{\infty} \frac{\lambda^n}{n!} \qquad r=k,k+1,\ldots;\ \lambda > 0 \tag{3.6.4}$$

$$b_n = e^{-\mu} \frac{\mu^n}{n!} \qquad n=0,1,\ldots;\ \mu > 0\ . \tag{3.6.5}$$

Then, for a_r and b_n as defined above, the corresponding c_n is going to be

$$c_n = \sum_{r=k}^{n} a_r b_{n-r} = \frac{\sum_{r=k}^{n} \frac{\lambda^r}{r!} e^{-\mu} \frac{\mu^{n-r}}{(n-r)!}}{\sum_{n=k}^{\infty} \frac{\lambda^n}{n!}}$$

i.e.

$$c_n = \frac{e^{-\mu} \sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r}}{n! \sum_{n=k}^{\infty} \frac{\lambda^n}{n!}} \qquad n=k,k+1,\ldots\ . \tag{3.6.6}$$

Consequently,

$$\frac{a_r b_{n-r}}{c_n} = \frac{\binom{n}{r} p^r q^{n-r}}{\sum_{r=k}^{n} \binom{n}{r} p^r q^{n-r}} \qquad r=k,k+1,\ldots$$

with $p = \frac{\lambda}{\lambda+\mu}$ $q = 1-p$. Hence, with the sequences a_r, b_n defined as in (3.6.4) and (3.6.5) the truncated binomial can be written in the form $\frac{a_r b_{n-r}}{c_n}$.

From theorem (3.6.1) we know that the R-R condition, as given by (3.5.3), holds iff

$$P_n = P_k \frac{c_n}{c_k} \theta^{n-k} \quad \text{for some } \theta > 0.$$

i.e. (by making use of (3.6.6)) iff

$$P_n = P_k \frac{\sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r} / n!}{\binom{k}{k} \lambda^k / k!} \theta^{n-k}$$

and since

$$\frac{1}{P_k} = \frac{\sum_{n=k}^{\infty} \sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r} / n!}{\lambda^k / k!} \theta^{n-k},$$

iff

$$P_n = \frac{\sum_{r=k}^{n} \binom{n}{r} (\lambda\theta)^r (\mu\theta)^{n-r}}{n! \sum_{n=k}^{\infty} \sum_{r=k}^{n} \binom{n}{r} (\lambda\theta)^r (\mu\theta)^{n-r} / n!} \tag{3.6.7}$$

But

$$\sum_{n=k}^{\infty} \sum_{r=k}^{n} \binom{n}{r} (\lambda\theta)^r (\mu\theta)^{n-r} / n! = \sum_{r=k}^{\infty} \sum_{n=r}^{\infty} \binom{n}{r} (\lambda\theta)^r (\mu\theta)^{n-r} / n!$$

$$= \sum_{r=k}^{\infty} \sum_{n=0}^{\infty} \frac{(\lambda\theta)^r (\mu\theta)^n}{r! n!} = e^{\mu\theta} \sum_{r=k}^{\infty} \frac{(\lambda\theta)^r}{r!}. \tag{3.6.8}$$

Hence (using (3.6.7) and (3.6.8)) we find that (3.5.3) is valid iff P_n is of the form (3.6.3).

Note

It is interesting to point out here, that the truncated Binomial is not the only distribution of the form (3.6.1) for which the previous result is valid. In other words, if we assume that P_n corresponds to the distribution which is the convolution of a Poisson with a truncated Poisson as it is given in (3.6.3) and $P(Y=r|X=n)$ is of the form $\frac{a_r b_{n-r}}{c_n}$ $r=k,k+1,\ldots$ then, the R-R condition (3.5.3) does not imply uniquely that the distribution of $Y|X$ is truncated Binomial. In fact from theorem 3.6.1 we know that the R-R condition is equivalent to

$$c_n = c_k \frac{P_n}{P_k} \theta^{-n+k} \quad n=k,k+1,\ldots \quad \text{for some } \theta > 0$$

which in this particular case implies that

$$c_n = c_k \frac{\sum_{r=k}^{n} \binom{n}{r} \lambda^r \mu^{n-r}/n!}{\lambda^k/k!} \theta^{-n+k} \quad n=k,k+1,\ldots \tag{3.6.9}$$

It can be checked easily that a_n, b_n can be of the form

$$a_r = a_k \frac{\lambda^r}{r!} \quad r=k,k+1,\ldots, \quad \lambda > 0, \; a_k \text{ constant} \tag{3.6.10}$$

$$b_n = b_0 \frac{\mu^n}{n!}, \; n=0,1,\ldots, \quad \mu > 0, \; b_0 \text{ constant.} \tag{3.6.11}$$

However (3.6.10) and (3.6.11) are not the only forms that a_n, b_n respectively can have in order that their convolution is of the same form as (3.6.9).

Take for example the following sequences

$$a'_r = \frac{e^{-\mu} \sum_{m=k}^{r} \binom{r}{m} \lambda^m \mu^{r-m}}{r! \sum_{n=k}^{\infty} \lambda^n/n!} \quad r=k,k+1,\ldots. \tag{3.6.12}$$

$$b'_n = e^{-\mu} \frac{\mu^n}{n!} \quad n=0,1,\ldots \tag{3.6.13}$$

i.e. take a'_n to be convolution of a Poisson with a truncated Poisson, and b'_n to be Poisson. Then,

$$\begin{aligned}\{a'_n\} * \{b'_n\} &\sim (\text{Poisson} * \text{truncated Poisson}) * \text{Poisson} \\ &\sim (\text{Poisson} * \text{Poisson}) * \text{truncated Poisson} \\ &\sim \text{Poisson} * \text{Truncated Poisson.}\end{aligned}$$

This means that we can find another pair of sequences, whose convolution is of the form (3.6.9). Hence the decomposition of (3.6.9) is not unique, and hence the truncated Binomial is not the only distribution of the form (3.6.1) satisfying our conditions.

The previous note, will help us to give an answer that arises out of a research note published by Moran (1952) and related to a characteristic property of the Poisson distribution. His result goes as follows (in our notation).

Let Y and Z be independent random variables taking non-negative integral values and let X = Y+Z. Suppose that there exists at least one *integer i so that*

$$P(Y=i) > 0, \quad P(Z=i) > 0.$$

Then, Y and Z are individually distributed in Poisson probability laws iff $P(Y|X)$ is binomial. It is natural now to ask whether a similar property holds in the case where $P(Y|X)$ is truncated binomial at the point k-1. The question can be formed as follows.

<u>Problem 3.6.1</u>

Let Y and Z be independent random variables, with Z taking non-negative values and Y taking values greater than or equal to k, $k \geq 0$ integer, and let X = Y+Z. Suppose that

$$P(Y=i) > 0, \quad P(Z=j) > 0$$

for at least one integer i, $i > k$ and one integer j. Then, is the condition that Y is Poisson truncated at k-1 and Z Poisson necessary and sufficient for $P(Y|X)$ to be binomial truncated at k-1? Evidently, the "sufficient" part is a side result of corollary 3.6.1. As far as the "necessary" part is concerned, the answer is negative. This is so, because, in corollary 3.6.1

we have shown that if $P(Y=r|X=n)$ is truncated binomial, then the R-R condition implies that X is convolution of a Poisson with a truncated Poisson. But as it was shown in the note, if the R-R condition holds then the truncated Poisson for Y and the Poisson for Z are not the only distributions for which X = Y+Z is Poisson convoluted with a truncated Poisson. Since the R-R condition is a special case of Y and Z being independent, the argument is established.

<u>Corollary 3.6.2</u> (Characterization of the Distribution which is the Convolution of a Negative Binomial with a Truncated Negative Binomial.)

Consider the r.v's X and Y as in theorem 3.6.1. Suppose that the conditional distribution of Y on X is Negative Hypergeometric, truncated at k-1, i.e.

$$P(Y=r|X=n) = \frac{\binom{-m}{r}\binom{-\rho}{n-r}}{\sum_{r=k}^{n}\binom{-m}{r}\binom{-\rho}{n-r}} \qquad r=k,k+1,\ldots \quad r \leq n. \tag{3.6.14}$$

Then the R-R condition (3.5.3) holds iff P_n is the convolution of a Negative Binomial (p,m) with a truncated Negative Binomial (p,ρ) at the point k-1, i.e. iff

$$P_n = \frac{\sum_{r=k}^{n}\binom{-m}{r}\binom{-\rho}{n-r}(-q)^n(1-q)^{\rho}}{\sum_{r=0}^{\infty}\binom{-m}{r}(-q)^r} \qquad n=k,k+1,\ldots \tag{3.6.15}$$

<u>Proof</u> Let

$$a_r = \frac{\binom{m+r-1}{r}q^r}{\sum_{r=k}^{\infty}\binom{m+r-1}{r}q^r} \qquad r=k,k+1,\ldots \tag{3.6.16}$$

and

$$b_n = \binom{\rho+n-1}{n} q^n \quad n=0,1\ldots \tag{3.6.17}$$

Then,

$$c_n = \frac{\sum_{r=k}^{n} \binom{m+r-1}{r} \binom{\rho+n-r-1}{n-r} q^n}{\sum_{r=k}^{\infty} \binom{m+r-1}{r} q^r} \quad n=k,k+1,\ldots \tag{3.6.18}$$

Clearly

$$\frac{a_r b_{n-r}}{c_n} = \frac{\binom{-m}{r} \binom{-\rho}{n-r}}{\sum_{r=k}^{\infty} \binom{-m}{r} \binom{-\rho}{n-r}} \quad r=k,k+1,\ldots$$

i.e. the above defined sequences a_r, b_n can be used to express (3.6.14) in the form required by theorem 3.6.1. Accordingly the R-R condition (3.5.3) will be satisfied iff

$$P_n = P_k \frac{c_n}{c_k} \theta^{n-k} \quad \text{for a suitable } \theta > 0$$

or equivalently (using 3.6.18) iff

$$P_n = P_k \frac{\sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-1)^n \theta^{n-k}}{\binom{-m}{k} \binom{-\rho}{n-k} (-1)^k} .$$

But because $\sum_{n=k}^{\infty} P_n = 1$,

$$P_k^{-1} = \frac{\sum_{n=k}^{\infty} \sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-\theta)^{n-k}}{\binom{-m}{k} \binom{-\rho}{n-k}} = \frac{\sum_{r=k}^{\infty} \sum_{n=r}^{\infty} \binom{-m}{r} \binom{-\rho}{n-r} (-\theta)^{n-k}}{\binom{-m}{k} \binom{-\rho}{n-k}}$$

$$= \frac{\sum_{r=k}^{\infty} \binom{-m}{r} (-\theta)^{r-k} \sum_{n=0}^{\infty} \binom{-\rho}{n} (-\theta)^{n}}{\binom{-m}{k} \binom{-\rho}{n-k}} = \frac{\sum_{r=k}^{\infty} \binom{-m}{r} (-\theta)^{r-k} (1-\theta)^{-\rho}}{\binom{-m}{k} \binom{-\rho}{n-k}} .$$

Hence (3.5.3) holds iff

$$P_n = \frac{\sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-\theta)^{n-k}}{\sum_{r=k}^{\infty} \binom{-m}{r} (-\theta)^{r-k} (1-\theta)^{-\rho}}$$

i.e.

$$P_n = \frac{\sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-q)^{n} (1-q)^{\rho}}{\sum_{r=k}^{\infty} \binom{-m}{r} (-q)^{r}}$$

for $q = \theta$ with $0 < \theta < 1$.

<u>Note</u>

Here again, as in the case of the truncated Binomial, if we assume that P_n is the convolution of a Negative Binomial with a Truncated Negative Binomial as given by (3.6.15) and $P(Y=r|X=n)$ is of the form $\frac{a_r b_{n-r}}{c_n}$ $r=k,k+1,\ldots$ we will get that the R-R condition (3.5.3) does not imply uniquely that the distribution of $Y|X$ is truncated Negative Hyper-geometric. This can be seen by making use of the fact that under the circumstances mentioned, from theorem 3.6.1 we have that the R-R condition is valid iff $c_n = c_k \frac{P_n}{P_k} \theta^{-n+k}$, which for P_n as in (3.6.15) will eventually become (following the same steps as in the Note of Corollary 3.6.1).

$$c_n = \sum_{r=k}^{n} a_r b_{n-r} = \frac{\sum_{r=k}^{n} \binom{-m}{r} \binom{-\rho}{n-r} (-q/\theta)^{n} (1-q/\theta)^{\rho}}{\sum_{r=0}^{\infty} \binom{-m}{r} (-q/\theta)^{r}} \qquad (3.6.19)$$

for some $0 < q/\theta < 1$ and a_r, b_n probability distributions.

It is evident by observing (3.6.16), (3.6.17) and (3.6.18) that the truncated Negative Binomial as a_r and the Negative Binomial as b_n is a solution of (3.6.19). However, this is not again the only solution. For, consider the example

$$a'_r = \frac{\sum_{i=k}^{r} \binom{-m}{i} \binom{-\rho}{r-i} p^{\rho} (-q)^r}{\sum_{i=k}^{\infty} \binom{-m}{i} (-q)^i} \qquad r=k,k+1,\ldots \tag{3.6.20}$$

i.e. convolution of a truncated Negative Binomial with a Negative Binomial and

$$b'_n = \binom{-\rho}{n} p^{\rho} (-q)^n \qquad n=0,1,\ldots \tag{3.6.21}$$

i.e. a Negative Binomial.

Then,

$$\begin{aligned}\{c'_n\} = \{a'_n\}*\{b'_n\} &\sim (\text{Negative Binomial} * \text{truncated Neg. Bin.}) * \text{Neg. Bin.} \\ &\sim \text{Neg. Bin.} * \text{truncated Neg. Bin.}\end{aligned} \tag{3.6.22}$$

Clearly $\{c'_n\}$ is of the same form as c_n in (3.6.19).

But the distribution $\frac{a'_r b'_{n-r}}{c'_n}$ with a'_r, b'_n, c'_n given by (3.6.20), (3.6.21) and (3.6.22) is not truncated Negative Hypergeometric. This means that there exists at least one distribution $Y|X$ other than the truncated Negative Binomial for which Corollary 3.6.2 is true.

<u>Remark 1</u> Patil and Seshadri (1964) show that if Y, Z are independent then $Y|X$ is Negative Hypergeometric, (where $X = Y+Z$), iff Y and Z are Negative Binomials. The previous note shows that for $Y|X$ truncated Negative Hypergeometric, a corresponding result with Y~truncated Negative

Binomial and Z~Negative Binomial is true only so far as the "if" part is concerned.

It may also be noted, that for reasons given in Note 2 of Corollary 3.3.2 X~Negative Binomial does not imply that $Y|X$ is Negative Hypergeometric in Patil and Seshadri's set up.

Remark 2 It is interesting to observe, that Corollaries 3.6.1, 3.6.2 can also be considered as special cases of the following theorem, which in fact is an extension of theorem 3.5.1 and a variant of theorem 3.6.1.

Theorem 3.6.2

Let $\{(a_n, b_n): n=0,1,\ldots\}$ be a sequence of vectors of non-negative real numbers such that $a_n > 0$ for $n \geq k$ and $b_0, b_1 > 0$, $b_n \geq 0$, $n=2,3,\ldots$ Let $\{c_n\}$ be the convolution of $\{a_n\}$ and $\{b_n\}$ (observe that $c_n > 0$, $n \geq k$).

Let (X,Y) be a vector of non-negative integer-valued r.v.'s such that:

$$P(X=n) = P_n \text{ with } P_k < 1.$$

Also, whenever $P_n > 0$

$$P(Y=r|X=n) = \frac{a_r b_{n-r}}{c_n} \qquad \begin{array}{l} r=0,1,\ldots n \\ n=k,k+1,\ldots \end{array} \tag{3.6.23}$$

Then,

$$P(Y=r|Y \geq k) = P(Y=r|X=Y) \qquad r=k,k+1,\ldots. \tag{3.6.24}$$

iff

$$\frac{P_n}{c_n} = \frac{P_k}{c_k}\,\theta^{n-k} \qquad \begin{array}{l} n=k,k+1,\ldots \\ \text{for some } \theta > 0. \end{array} \tag{3.6.25}$$

Proof

We have been given that X-k is non-negative integer-valued random variable. Further, it follows that conditional on $Y-k \geq 0$, the r.v. Y-k is non-negative integer-valued.

If we define

$$c_n^{(k)} = \sum_{r=k}^{n} a_r b_{n-r} \qquad (3.6.26)$$

we will have

$$c_{n+k}^{(k)} = \sum_{r=k}^{n+k} a_r b_{n+k-r} = \sum_{r=0}^{n} a_{r+k} b_{n-r} . \qquad (3.6.27)$$

It is

$$P(X\text{-}k=n|Y\text{-}k \geq 0) = \frac{1}{P(Y \geq k)} P_{n+k} \frac{c_{n+k}^{(k)}}{c_{n+k}} \qquad n=0,1,... \qquad (3.6.28)$$

and

$$P(Y\text{-}k=r|X\text{-}k=n, Y\text{-}k \geq 0) = \frac{a_{r+k} b_{n-r}}{c_{n+k}^{(k)}} \qquad \begin{matrix} r=0,1,...,n \\ n=0,1,... \end{matrix} \qquad (3.6.29)$$

It also follows that (3.6.24) is equivalent to

$$P(Y\text{-}k=r|Y\text{-}k \geq 0) = P(Y\text{-}k=r|X\text{-}k=Y\text{-}k), \quad r=0,1,... \qquad (3.6.30)$$

If we now apply the result of theorem 3.1.1 to random variables X-k and Y-k and consider P_n, n=0,1,... to be given by (3.6.28), the result follows.

Note

Corollaries 3.6.1 and 3.6.2 can be derived from theorem 3.6.2 by considering the sequence a_n as a sequence with two parts $a_n = 0$ for n=0,1,...,k-1 and as in (3.6.4) and (3.6.16), respectively, for $n \geq k$.