isid/ms/2008/08
November 11, 2008
http://www.isid.ac.in/~statmath/eprints

Characterization of discrete distributions by conditional variance

K. K. Sudheesh
N. Unnikrishnan Nair

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

CHARACTERIZATION OF DISCRETE DISTRIBUTIONS BY CONDITIONAL VARIANCE

K. K. Sudheesh*,† and N. Unnikrishnan Nair**

*Indian Statistical Institute, Delhi Centre,

New Delhi-16, India.

**Cochin University of Science and Technology,

Cochin-22, India.

Abstract. In this paper we discuss characterization of a class of discrete distributions by properties of conditional variance. These properties include relationship between variance residual life, mean life function and the failure rate.

Keywords: Characterization, conditional variance, failure rate.

Classification code: 60 K 10, 62 N 05

1. Introduction

In studying the lifetime of a device or organism, the concepts of remaining life based on the current age is effectively used in reliability and survival analysis to infer properties of the underlying life distribution. In this context characterizations of life distributions based on the properties of various functions of remaining life such as its mean, median, percentiles variance etc have been extensively studied in literature. Of these, the role of variance residual life in determining the life distribution has been discussed for continuous models by various authors like Gupta (1987), Gupta et al. (1987), Hitha and Nair (1989), Gupta and Kirmani (2000,2004) and El- Arishi (2005). It remains an open problem to find similar characterizations for discrete distributions. Accordingly in this paper we discuss the general conditions under which discrete distributions can be characterized by the properties of conditional variance and specialize our results for some well known families of distributions.

†*Corresponding author*: E-mail:skkattu@yahoo.co.in, Fax:91-11-41493981, Phone:91-11-41493961.

2. Characterizations

Let X be a discrete random variable defined on the set of non-negative integers with distribution function $F(x)$ and probability mass function $f(x)$. The failure rate $k(x)$ and reversed failure rate $\lambda(x)$ of X are

$$k(x) = f(x)/R(x) \text{ and } \lambda(x) = f(x)/F(x)$$

where $R(x) = P(X \geq x)$. Defining $\mathcal{B}$ to be the class of real valued functions of x we denote the conditional expectations of $h(x) \in \mathcal{B}$, satisfying $E(h^2(X)) < \infty$, as

$$m(x) = E(h(X)|X > x),\ r(x) = E(h(X)|X \leq x)$$

and the mean and variance by

$$\mu = E(h(X),\ \sigma^2 = V(h(X))).$$

With these notations, Nair and Sudheesh (2008) established that probability mass function of X has the form

$$f(x+1)/f(x) = \sigma g(x)/(\sigma g(x+1) - \mu + h(x+1)) \tag{2.1}$$

with $\sigma g(0) = \mu - h(0)$, for some $g(x) \in \mathcal{B}$ if and only if

$$r(x) = \mu - \sigma g(x)\lambda(x) \tag{2.2}$$

or equivalently

$$m(x) = \mu + \sigma g(x)k(x)/(1 - k(x)). \tag{2.3}$$

For many of the discrete distributions neither of the functions r, m, λ and k have simple tractable forms and hence (2.2) and (2.3) are simple relationships that could be used for identifying the appropriate model through the form of $g(x)$ function which is unique for a particular distribution, given a choice of $h(x)$. Our first result is an identity connecting the conditional covariance of $h(x)$ and any $c(x) \in \mathcal{B}$ with the above quantities. In the sequel we use Δ as the usual forward difference operator.

Theorem 1. Let $c(x)$ and $h(x)$ be functions in $\mathcal{B}$ satisfying

(i) $\Delta c(x) \neq 0$

(ii) $E(|\Delta c(X)|g(X)),\ E(c(X)h(X)),\ E(h^2(X))$ are all $< \infty$.

Then X has distribution (2.1) for some $g(x) \in \mathcal{B}$ if and only if for all $x = 0, 1, 2, ...$

$$Cov(c(X), h(X)|X > x) = \sigma E(\Delta c(X).g(X)|X > x) + (\mu - m(x))(a(x) - c(x+1)) \qquad (2.4)$$

where $a(x) = E(c(X)|X > x)$.

Proof: First, we note that

$$E(c(X)(h(X) - \mu)|X > x) = [R(x+1)]^{-1} \sum_{k=x+1}^{\infty} c(k)(h(k) - \mu)f(k)$$

$$= [R(x+1)]^{-1} \sum_{t=1}^{\infty} c(x+t)\Big[\sum_{k=x+t}^{\infty}(h(k) - \mu)f(k) - \sum_{k=x+t+1}^{\infty}(h(k) - \mu)f(k)\Big]$$

$$= [R(x+1)]^{-1}\Big[\sum_{t=x+1}^{\infty}\Delta c(t)\sum_{k=t+1}^{\infty}(h(k) - \mu)f(k)\Big] + c(x+1)(m(x) - \mu)$$

$$= [R(x+1)]^{-1}\Big[\sum_{t=x+1}^{\infty}\Delta c(t)\sum_{k=0}^{t}(\mu - h(k))f(k)\Big] + c(x+1)(m(x) - \mu), \qquad (2.5)$$

since $E(h(X) - \mu) = \sum_{k=0}^{\infty}(h(k) - \mu)f(k) = 0$.

Now, suppose that (2.1) is true. Then from (2.2)

$$\sigma f(x)g(x) = \sum_{k=0}^{x}(\mu - h(k))f(k)$$

and substituting this in (2.5)

$$E(c(X)(h(X) - \mu)|X > x) = [R(x+1)]^{-1}\sigma \sum_{t=x+1}^{\infty}\Delta c(t).g(t)f(t)$$

$$+c(x+1)(m(x) - \mu)$$

$$= \sigma E(\Delta c(X).g(X)|X > x) + c(x+1)(m(x) - \mu), \qquad (2.6)$$

or

$$E(c(X)h(X)|X > x) = \sigma E(\Delta c(X).g(X)|X > x) + c(x+1)(m(x) - \mu) + \mu a(x),$$

which leads to (2.4).

Conversely, if we assume (2.4) by retracing the above steps we have (2.5). Using (2.5) and (2.6), the following equation results.

$$\sum_{t=x+1}^{\infty} \Delta c(t) \sum_{k=0}^{t} (\mu - h(k)) f(k) = \sigma \sum_{t=x+1}^{\infty} \Delta c(t) g(t) f(t)$$

Changing x to $x-1$ in the last equation and subtracting

$$\sum_{k=0}^{x} (\mu - h(k)) f(k) = \sigma g(x) f(x), \tag{2.7}$$

since by assumption $\Delta c(x) \neq 0$. However, (2.7) is the same as (2.2) and hence (2.1) holds.

Remark 1. Since $c(x)$ is an arbitrary function in $\mathcal{B}$, we can set $c(x) = h(x)$ in (2.4) so that

$$V(h(X)|X > x) = \sigma E(\Delta h(X).g(X)|X > x) + (\mu - m(x)) + (m(x) - h(x+1)) \tag{2.8}$$

for all x,characterizes the distribution specified in (2.1). When $h(X) = X$, we have the identity for the variance residual life in terms of the mean life function

$$m_1(x) = E(X|X > x)$$

given as

$$V(X|X > x) = \sigma E(g(X)|X > x) + (\mu - m_1(x)) + (m_1(x) - x - 1)) \tag{2.9}$$

with $\mu = E(X)$ and $g(x)$ arising from (2.2) or (2.3), that characterizes

$$f(x+1)/f(x) = \sigma g(x)/(\sigma g(x+1) - \mu + h(x+1)), x = 0, 1, 2...$$

We give several examples at the end of this section of simple forms (2.9) that is specified to standard discrete distributions.

Remark 2. In terms of the failure rate $k(x)$ of X, we can rewrite (2.8) as

$$V(h(X)|X > x) = \sigma E(\Delta h(X).g(X)|X > x) + (h(x+1) - \mu)\sigma g(x)k(x)(1 - k(x))^{-1}$$

$$-\sigma^2 g^2(x) k^2(x)(1 - k(x))^{-2}$$

as the identity connecting the conditional variance and failure rate for all distributions of the form (2.1).

We now examine how the above characteristic properties assume forms when specialized to some well known families of distributions.

Theorem 2. The distribution of X belongs to the modified power series family with

$$f(x) = a(x)(u(\theta))^x/A(\theta), x = 0, 1, 2, ... \tag{2.10}$$

$a(x) > 0$, $u(\theta)$ and $A(\theta)$ are positive, finite and differentiable, if and only if

$$V(X|X > x) = \frac{u(\theta)}{u'(\theta)}\frac{\partial m_1(x)}{\partial\theta} \tag{2.11}$$

Proof: From Nair and Sudheesh (2008), for the family (2.10)

$$g(x) = -\frac{\mu}{\sigma}\frac{A(\theta)}{A'(\theta)}\frac{1}{f(x)}\frac{\partial F(x))}{\partial\theta} = -\frac{u(\theta)}{u'(\theta)}\frac{1}{\sigma f(x)}\frac{\partial F(x)}{\partial\theta}.$$

Hence when $h(x) = x$

$$\sigma E(\Delta X.g(X)|X > x) = \frac{u(\theta)}{u'(\theta)R(x+1)}\sum_{t=x+1}^{\infty}(t-(t+1))\frac{\partial F(t)}{\partial\theta}$$

$$= \frac{u(\theta)}{u'(\theta)R(x+1)}\frac{\partial}{\partial\theta}[(x+1)F(x+1) + \sum_{t=x+2}^{\infty} tf(t)]$$

$$= \frac{u(\theta)}{u'(\theta)R(x+1)}\frac{\partial}{\partial\theta}[(x+1)(F(x+1) - f(x+1)) + \sum_{t=x+1}^{\infty} tf(t)]$$

$$= \frac{u(\theta)}{u'(\theta)R(x+1)}\frac{\partial}{\partial\theta}[m_1(x)R(x+1) + (x+1)F(x)]. \tag{2.12}$$

Now,

$$\frac{\partial}{\partial\theta}[m_1(x)R(x+1) + (x+1)F(x)] = R(x+1)\frac{\partial m_1(x)}{\partial\theta} + m_1(x)\frac{\partial R(x+1)}{\partial\theta} + (x+1)\frac{\partial F(x)}{\partial\theta}$$

$$= R(x+1)\frac{\partial m_1(x)}{\partial\theta} + ((x+1) - m_1(x))\frac{\partial F(x)}{\partial\theta}. \tag{2.13}$$

Using (2.13) in (2.12) and noting the expression for $g(x)$ above,

$$\sigma E(g(X)|X > x) = \frac{u(\theta)}{u'(\theta)}\frac{\partial m_1(x)}{\partial\theta} + (m_1(x) - \mu)(m_1(x) - x - 1).$$

The final form of $V(X|X > x)$ as stated in the Theorem is recovered from (2.9). The converse part is obtained by retracing the above steps to arrive at $g(x)$ and hence $f(x)$.

Remark 3. The results for the generalized power series family is obtained by setting $u(\theta) = \theta$.

Theorem 3. X belongs to the Ord family of distributions defined by

$$\frac{f(x+1) - f(x)}{f(x)} = \frac{-(x+d)}{a_0 + a_1 x + a_2 x^2} \tag{2.14}$$

if and only if

$$V(X|X > x) = E((b_0 + b_1 X + b_2 X^2)|X > x) + (\mu - m_1(x))(m_1(x) - x - 1)$$

where

$$b_0 = \mu + \frac{a_0 - a_1 + a_2}{1 - 2a_2}, b_1 = \frac{a_1 - 1}{1 - 2a_2}, b_2 = \frac{a_2}{1 - 2a_2}.$$

Proof: Writing (2.14) in the form

$$\frac{f(x+1)}{f(x)} = \frac{(a_0 - d) + (a_1 - 1)x + a_2 x^2}{a_0 + a_1 x + a_2 x^2} = \frac{\sigma g(x)}{\sigma g(x+1) - \mu + (x+1)},$$

$g(x)$ must be a quadratic of the form $b_0 + b_1 x + b_2 x^2$. Substituting into the last equation and equating like coefficients, the value of b_0, b_1 and b_2 are obtained as in the Theorem. The rest of the proof is evident from Remark 2.1.

As a particular case of the Ord family, we have the Katz family which is important in many applications in its own right.

Theorem 4. X belongs to Katz family

$$\frac{f(x+1)}{f(x)} = \frac{\alpha + \beta x}{1 + x},$$

if and only if

$$V(X|X > x) = (1 - \beta)^{-1}(\alpha + \beta m_1(x)) + (\mu - m_1(x))(m_1(x) - x - 1).$$

The last result is a straight forward application of

$$\sigma g(x) = (1 - \beta)^{-1}(\alpha + \beta x)$$

for the Katz family, in (2.9).

We now illustrate our results for some specific distributions belonging to the above families.

Example 1. For the binomial distribution with parameters n and p, $\sigma g(x) = p(n-x)$ and hence

$$V(X|X>x) = m_1(x)) + (x+\mu - m_1(x)) - \mu x$$

and also

$$V(X|X>x) = np(1-p) + (x+1)(1-p)(x+1-p(n+1))k(x+1) - (x+1)^2k^2(x+1).$$

Example 2. In the Poisson case

$$f(x) = e^{-\lambda}\lambda^x/x!,$$

$\sigma g(x) = \lambda$. Thus from (2.9)

$$V(X|X>x) = m_1(x))(\mu - m_1(x)) - \mu x$$

$$= \lambda + (x+1)(x+1-\lambda)k(x+1) - (x+1)^2k^2(x+1).$$

Remark 4. The expressions for the binomial and Poisson distributions are identical with those of El-Arishi (2005) obtained using a different approach. Equation (2.9) subsumes El-Arishi formulas as special cases.

Example 3. The negative binomial distribution

$$f(x) = \binom{n+x-1}{n-1} p^n(1-p)^x, x = 0,1,2...$$

as a member of Ord family provides $\sigma g(x) = p^{-1}(1-p)(x+n)$ giving

$$V(X|X>x) = m_1(x))(2\mu - m_1(x) + x + 1) - \mu x$$

$$= p^{-1}\mu + p^{-1}(2(x+1) - \mu)k(x+1) - (x+1)^2k^2(x+1).$$

In particular for the geometric law

$$m_1(x) = \mu + x + 1$$

$$V(X|X>x) = V(X)$$

and further in terms of the mean function

$$V(X|X>x) = (m_1(x) - x)(m_1(x) - x - 1).$$

The geometric distribution, Waring distribution specified by

$$f(x) = \frac{(a-b)(b)_x}{(a)_{x+1}}, (b)_x = b(b-1)...(b-x+1)$$

and the negative hypergeometric with

$$f(x) = \binom{-1}{x}\binom{-a}{n-x}\Big/\binom{-1-a}{n}, x = 0, 1, 2...$$

share a common characteristic property

$$V(X|X > x) = C(m_1(x) - x)(m_1(x) - x - 1).$$

where $C = 1(< 1; > 1)$ for the geometric(negative hypergeometric; Waring).
Notice further that in this case, the variance residual life function is quadratic function of the mean residual life. More examples from modified power series family can be constructed using $g(x)$ values given in Nair and Sudheesh (2008).

References

1. El- Arishi, S (2005). A conditional variance characterization of some discrete probability distributions, *Statist. Papers,* 46, 31- 45.

2. Gupta, R.C (1987). On the monotonic properties of residual variance and their applications in reliability, *J. Statist. Plann. Inf.*, 16, 329-335.

3. Gupta, R.C., Kirmani, S.N.U.A. and Launer, R.L (1987). On life distributions having monotone residual variance, *Probab. Engrg. Inform. Sci.*, 1, 299- 307.

4. Gupta, R.C. and Kirmani, S.N.U.A (2000). Residual coefficient of variation and some characterization results, *J. Statist. Plann. Inf.*, 91, 23-31.

5. Gupta, R.C. and Kirmani, S.N.U.A (2004). Some characterization of distributions by functions of failure rates and mean residual life, *Comm. Statist. Theory Methods*, 33, 3115-3131.

6. Hitha, N. and Nair, N.U (1989). Characterization of some models by properties of residual life, *Calcutta Statist. Assoc. Bull.*, 38, 219-225.

7. Nair, N. U and Sudheesh. K. K (2008), Some Results on Lower Variance Bound Useful in Reliability and Estimation, *Ann. Inst. Stat. Math.*, 60, 591- 603.