Statistics & Probability Letters 45 (1999) 49–53

STATISTICS & PROBABILITY LETTERS

www.elsevier.nl/locate/stapro

A characterization of the negative binomial distribution via α-monotonicity

Theofanis Sapatinas

Institute of Mathematics and Statistics, University of Kent at Canterbury, Canterbury, Kent CT2 7NF, UK

Received November 1998; received in revised form January 1999

Abstract

We obtain a characterization of the negative binomial distribution by using arguments based on power-series families of distributions for α-monotone random variables. © 1999 Elsevier Science B.V. All rights reserved

Keywords: Power-series families of distributions; α-unimodality; α-monotonicity; Negative binomial distribution

1. Introduction

The notions of unimodality exist for both non-lattice as well as lattice distributions, each one having its own interpretation. Let X be a real-valued random variable (rv) and let F be its distribution function. The rv X (or F) is said to be unimodal about 0 (or with mode at 0) if F is convex on $(-\infty, 0)$ and concave on $(0, \infty)$. Olshen and Savage (1970) generalized the concept of unimodality on $\mathbb{R} = (-\infty, \infty)$ to α-unimodality and proved that a real-valued rv X is α-unimodal about 0, where $\alpha > 0$, if and only if

$$X \stackrel{\mathrm{d}}{=} UZ, \tag{1}$$

where $U \sim$ Beta$(\alpha, 1)$ and independent of the rv Z. (The case $\alpha = 1$ is the famous characterization of unimodal about 0 distributions on $\mathbb{R}$ due to Khintchine (1938).)

Discrete unimodal distributions have a different structure from that of unimodal distributions on $\mathbb{R}$. Steutel (1988) gave a definition analogous to Eq. (1), relative to discrete distributions. Restricting to distributions on $N_0 = \{0, 1, 2, \ldots\}$, he defined a rv X to be α-monotone (discrete α-unimodal on N_0), where $\alpha > 0$, if:

$$X \stackrel{\mathrm{d}}{=} U \otimes Z, \tag{2}$$

where $U \sim$ Beta$(\alpha, 1)$ and independent of the non-negative integer-valued rv Z; the multiplication $u \otimes X$ is defined (see Steutel and van Harn (1979)) by $\sum_{i=0}^{X} B_i$, where $\{B_i\}$ is a sequence of independent Bernoulli rv's with success probability u, and X is a non-negative integer-valued rv independent of $\{B_i\}$. For a detailed account on unimodality of probability measures, we refer to the recent monograph of Bertin et al. (1997).

E-mail address: t.sapatinas@ukc.ac.uk (T. Sapatinas)

PII: S0167-7152(99)00041-3

Over the last decade, characterization results based on the product forms (1), (2) have arisen. The emphasis in most of these results is not on unimodality per se, but rather the distribution of Z is chosen to be closely related to that of X, and to be stochastically larger than X. For example, sums or the maximum of independent copies of X have been considered. The goal is to identify distributions which are left invariant under a random contraction of Z – see, e.g. van Harn and Steutel (1993), Sapatinas (1995), Bertin et al. (1997) and references there in. Incidentally, van Harn and Steutel (1993) have also studied the case when the multiplication $\otimes$ is generalized by a lattice-preserving operation whose definition involves a subcritical Markov branching process. This was extended to the general case where U is any rv with support on $(0,1)$ by Pakes (1995).

In this note, we look at a characterization in which the distribution of Z is not chosen to be related to that of X in the ways mentioned above. Rather we take X and Z to share a structural property – that the distributions of X and Z belong to discrete exponential families of distributions. Specifically, we try to identify discrete distributions comprising a power-series (PS) family which are α-monotone (in the sense (2)) and with Z having a distribution belonging to a modified power-series (MPS) family. (PS and MPS families are the discrete analogues of natural exponential and exponential families respectively, see Johnson et al. (1993), pp. 70, 74.) We will show, subject to some positivity assumptions, that X and Z necessarily have negative binomial distributions.

2. The characterization

Let X and Z be non-negative integer-valued rv's. We assume that X has a distribution belonging to a PS family depending on a parameter θ with probability mass function (pmf) given by

$$p_n(\theta)=\frac{b_n\theta^n}{B(\theta)},\quad b_0>0,\ b_i>0\ \text{ for some } i\geqslant 2, \tag{3}$$

where $B(\theta)$ is the normalizing function and θ takes values in the interval $[0,\rho)$, $0<\rho\leqslant\infty$, where ρ is the radius of convergence of $B(\theta)$.

Theorem. *Assume that X and Z as defined above satisfy Eq.* (2) *for all θ. Then, Z has a distribution belonging to a MPS family with* pmf *given by*

$$g_n(\theta)=\frac{c_n(a(\theta))^n}{C(a(\theta))},\quad c_0>0, \tag{4}$$

where $C(a(\theta))$ is the normalizing constant, $a(\theta)$ is a non-negative function, differentiable with respect to θ and is not a constant, if and only if

$$X\sim \text{Negative binomial } (\alpha,\cdot).$$

(*Furthermore, if* $X\sim$ Negative binomial $(\alpha,\cdot)$ *then* $Z\sim$ Negative binomial $(\alpha+1,\cdot)$ *with* '$\cdot$' *referring to the same parameter.*)

Proof. The 'if' part of the assertion can be easily verified by substitution. We shall establish the 'only if' part of the assertion. Let P and G be the probability generating functions (pgfs) of X and Z, respectively. In terms of pgfs, Eq. (2) is equivalent to stating that

$$P(t)=\alpha\int_0^1 G(1-u(1-t))u^{\alpha-1}\,\mathrm{d}u,\quad t\in[-1,1],$$

which can be written as

$$G(t)=P(t)-\frac{(1-t)}{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\{P(t)\}. \tag{5}$$

On using the assumptions of the theorem, Eq. (5) is seen to be equivalent to

$$\sum_{n=0}^{\infty} c_n(a(\theta)t)^n = h(\theta)\left\{\sum_{n=0}^{\infty} b_n(\theta t)^n - \frac{1}{\alpha}\sum_{n=0}^{\infty}(n+1)b_{n+1}\theta^{n+1}t^n + \frac{1}{\alpha}\sum_{n=0}^{\infty} nb_n(\theta t)^n\right\}, \tag{6}$$

where $h(\theta)=C(a(\theta))/B(\theta)$. Putting $t=0$ in Eq. (6), we get $h(\theta)=c_0\alpha/(b_0\alpha-b_1\theta)$. Equating coefficients of t^n in Eq. (6), we get

$$c_n(a(\theta))^n = \frac{h(\theta)\theta^n}{\alpha}(b_n(n+\alpha)-b_{n+1}(n+1)\theta). \tag{7}$$

Observe that $a(0)=0$ and consider first the case $b_1=0$. Obviously, we have that

$$C(a(\theta))=c_0+c_1a(\theta)+c_2a(\theta)^2+\cdots\equiv(c_0/b_0)(b_0+b_2\theta^2\cdots). \tag{8}$$

It follows from Eq. (8) that $c_1=0$, and hence from Eq. (7) that $b_2=0$. Going back and forth between Eqs. (7) and (8) shows that only b_0 and c_0 are positive, i.e. $X=Z=0$ almost surely. Consider now the case $b_1>0$ and $b_2=0$. It follows from Eq. (7) that $b_3=0$, otherwise we get $c_2<0$. By arguing inductively from Eq. (7) we get that $a(\theta)\propto\theta h(\theta)$ and $c_n, b_n=0$ for $n\geqslant 2$. Finally consider the case $b_2>0$. It follows that all b_i's and c_i's are positive and

$$a(\theta)=\lambda\theta \quad \text{for some } \lambda>0,$$

$$c_n=\frac{c_0}{b_0}b_n\left(\frac{n}{\alpha}+1\right)\lambda^{-n}, \quad n\geqslant 0,$$

$$\frac{(n+1)b_{n+1}}{(n+\alpha)b_n}=\beta, \quad \text{for some } \beta>0,\ n\geqslant 0,$$

which entails

$$b_n=b_0\binom{-\alpha}{n}(-\beta)^n, \quad n\geqslant 0, \tag{9}$$

$$c_n=c_0\binom{-(\alpha+1)}{n}(-\beta)^n\lambda^{-n}, \quad n\geqslant 0. \tag{10}$$

Eqs. (9) and (10) imply that $X\sim$ Negative binomial $(\alpha,\cdot)$ and that $Z\sim$ Negative binomial $(\alpha+1,\cdot)$. This completes the proof of the 'only if' part and hence of the theorem.

Remark 1. On taking $\alpha=1$, a characterization of the geometric distribution is arrived at.

Remark 2. If the assumption that $b_i>0$ for some $i\geqslant 2$ given in Eq. (3) is relaxed, then we do not arrive at the characterization in question. The following counterexample shows that if we take $b_2=0$ then there are PS families related by Eq. (2) which are not negative binomial distributions.

Example. Let X be a Bernoulli rv with success parameter $\theta/(1+\theta)$ such that $0<\theta<\alpha$. Then:

$$P(t)=\left(\frac{1+\theta t}{1+\theta}\right), \quad t\in[-1,1]$$

and, hence, G given by Eq. (5) is equal to

$$G(t)=\left(\frac{(\alpha-\theta)+(\alpha+1)\theta t}{\alpha(1+\theta)}\right), \quad t\in[-1,1].$$

The above pgf is again the pgf of a Bernoulli rv with success parameter $(1+\alpha)\theta/[\alpha(1+\theta)]$ and, thus, our claim is valid.

Remark 3. (i) The above example can easily be extended to show that X can have both a PS distribution and be represented by Eq. (2) with Z failing to have a MPS distribution. To see this, let X be a binomial rv with pgf:

$$P(t)=\left(\frac{1+\theta t}{1+\theta}\right)^N, \quad t\in[-1,1],$$

where N is a positive integer. Let $R=N/\alpha$, and restrict θ so that $\theta<1/R$. Then Eq. (2) holds with Z having the pgf:

$$G(t)=\left(\frac{1+\theta t}{1+\theta}\right)^N\left(\frac{1-R\theta+(1+R)\theta t}{1+\theta t}\right), \quad t\in[-1,1],$$

which can be parametrized to comprise a MPS distribution only when $N=1$.

(ii) A further question is whether Z having a PS distribution implies that X has a MPS distribution. We answer in the negative. To see this, let:

$$G(t)=\exp\{-\theta(1-t)\}, \quad \theta>0, \quad t\in[-1,1]$$

and consider Eq. (2) with $\alpha=1$. Then:

$$P(t)=\frac{1-\exp\{-\theta(1-t)\}}{\theta(1-t)}, \quad t\in[-1,1],$$

which generates the upper tail sums of the Poisson (θ) distribution, normalized to be a distribution. This is not a MPS distribution.

Acknowledgements

The author would like to thank the referee whose suggestions led to improvement to the presentation of this note, and for providing Remark 3. Helpful discussions with D.N. Shanbhag on an early draft of this note are also acknowledged.

References

Bertin, E.M.J., Cuculescu, I., Theodorescu, R., 1997. Unimodality of Probability Measures. Kluwer Academic Publishers, Dordrecht.
Johnson, N.L., Kotz, S., Kemp, A.W., 1993. Univariate Discrete Distributions, 2nd ed. Wiley, New York.
Khintchine, A.Y., 1938. On unimodal distributions. Izv. Nauchno-Issled. Inst. Mat. Mekh. Towsk. Gos., Univ. 2, 1–7 (in Russian).
Olshen, R.A., Savage, L.J., 1970. A generalized unimodality. J. Appl. Probab. 7, 21–34.

Pakes, A.G., 1995. Characterization of discrete laws via mixed sums and Markov branching processes. Stochastic Process. Appl. 55, 285–300.

Sapatinas, T., 1995. Characterizations of probability distributions based on discrete p-monotonicity. Statist. Probab. Lett. 24, 339–344.

Steutel, F.W., 1988. Note on discrete α-unimodality. Statistica Neerlandica 48, 137–140.

Steutel, F.W., van Harn, K., 1979. Discrete analogues of self-decomposability and stability. Ann. Probab. 7, 893–899.

van Harn, K., Steutel, F.W., 1993. Stability equations for processes with stationary independent increments using branching processes and Poisson mixtures. Stochastic Process. Appl. 45, 209–230.