INFINITE DIVISIBILITY OF INTEGER-VALUED RANDOM VARIABLES

By S. K. Katti

Florida State University, Tallahassee, Florida

1. Introduction. Proof of infinite divisibility (i.d. for brevity) of distributions such as the Poisson, the Normal and, the Gamma is available in text books. The proof hinges on the fact that if $\phi(t, \theta)$ is the characteristic function of the given distribution, then $\{\phi(t, \theta)\}^n$ can be written in the form $\phi(t, \theta^*)$. Clearly, one cannot expect this to happen in general situations. For example, the characteristic function $\phi(t, \theta) = \log\{1 - \theta \exp(it)\}/\log\{1 - \theta\}$ of the logarithmic distribution does not permit this form. The aim of this paper is to develop conditions which are necessary and sufficient if an integer valued random variable is to be i.d. and to illustrate their use with examples. The interesting part of these conditions is that they are explicit. Because of this explicitness, numerical methods can be used to give counter examples when a distribution is not i.d. and to gather inspiration to work out algebraic proofs if the numerical methods indicate that the distribution could be i.d.

2. Necessary and sufficient conditions. Let $g(z)$ be the probability generating function (pgf for brevity) of a random variable X taking values in $\{0, 1, 2, \cdots\}$ with probabilities P_i, $P_0 \neq 0$. If X is i.d., then for every integer n, we have $X = \sum_{i=1}^n X_i$ where X_i are independent and identically distributed. Clearly, X_i takes on the values $(0, 1, 2 \cdots)$. Denote the probabilities in the distribution of X_i by π_i and the pgf by $h(z)$. Then, $h(z) = \sum \pi_i z^i$ and

$$g(z) = \{h(z)\}^n. \tag{1}$$

On differentiating (1) with respect z, we get on rearranging the terms,

$$(\textstyle\sum_{i=0}^\infty (i+1)P_{i+1}^* z^i)(1 + \sum_{i=1}^\infty \pi_i^* z^i/n) = (1 + \sum_{i=1}^\infty P_i^* z^i)(\sum_{i=0}^\infty (i+1)\pi_{i+1}^* z^i), \tag{2}$$

where $P_i^* = P_i/P_0$ and $\pi_i^* = n\pi_i/\pi_0$. On ignoring the term $\sum_{i=1} \pi_i^* z^i/n$ in comparison with unity on the left-hand side of (2) we get

$$\textstyle\sum_{i=0}^\infty (i+1)P_{i+1}^* z^i = (1 + \sum_{i=1}^\infty P_i^* z^i)(\sum_{i=0}^\infty (i+1)\pi_{i+1}^* z^i). \tag{3}$$

On equating coefficients of z^i and solving for π_i^*, we get

$$\pi_i^* = (-1)^{i+1} \begin{vmatrix} P_1^* & 1 & 0 & \cdots & 0 & 0 \\ P_2^* & P_1^* & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \\ P_{i-1}^* & P_{i-2}^* & P_{i-3}^* & \cdots & P_1^* & 1 \\ iP_i^* & (i-1)P_{i-1}^* & (i-2)P_{i-2}^* & \cdots & 2P_2^* & P_1^* \end{vmatrix} \tag{4}$$

Received 14 December 1965; revised 26 August 1966.

or

$$\pi_i^* = iP_i^* - \sum_{j=1}^{i-1} P_j^* \pi_{i-j}^* . \tag{5}$$

A set of necessary conditions for the random variable X to be i.d. therefore is $\pi_i^* \geqq 0$ where π_i^* is given explicitly by (4) or through a recursion formula by (5). Sufficiency is clear from the representation in formula (7) or from the fact that the pgf $h(z)$ in (1) may be written as

$$h(z) = h(0) \exp [1/n \sum_{i=1}^{\infty} \pi_i^* z^i].$$

Corollary 1. *If we solve the equations given by* (3) *without ignoring the terms, we get the following solutions for* π_i^*, *denoted by* $\pi_i^{*(n)}$ *to distinguish them from the solution in formula* (4):

$$\pi_1^*(n) = (-1)^{i+1} \tag{6}$$

$$\cdot \begin{vmatrix} P_1^*(1 - 1/(i-1)n) & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \\ P_{i-2}^*(1-(i-2)/2n) & P_{i-3}^*(1-(i-3)/2n) & P_{i-4}^*(1-(i-4)/2n) & \cdots & 1 & 0 \\ P_i^*(1-(i-1)/n) & P_{i-2}^*(1-(i-2)/n) & P_{i-3}^*(1-(i-3)/n) & \cdots & P_1(1-1/n) & 1 \\ iP_i^* & (i-1)P_{i-1}^* & (i-2)P_{i-2}^* & \cdots & 2P_2^* & P_1^* \end{vmatrix}.$$

Since the conditions $\pi_i^* \geqq 0$ and $\pi_i^*(n) \geqq 0$ are both necessary and sufficient for the distribution of X to be i.d., we get the result that if the determinant on the right-hand side of (4) is nonnegative then the determinant on the right-hand side of (6) is nonnegative for every integer n.

Corollary 2. *We will prove here that if* $\pi_i^* > 0$ *for all* i, *then* $\pi_i^*(n) > \pi_i(n) > 0$ *for* $i \geqq 2$.

Denote the column vectors $(\pi_1^*(n), \pi_2^*(n), \cdots)'$ and $(\pi_1^*, \pi_2^*, \cdots)'$ by $\pi^*(n)$ and π^*, respectively. Then it can be shown that

$$\pi^*(n) = (1 + A/n + A^2/n^2 + \cdots)\pi^*, \tag{7}$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ \pi_1^* & 0 & 0 & 0 & \cdots \\ \pi_2^* & \pi_1^*/2 & 0 & 0 & \cdots \\ \pi_3^* & \pi_2^*/2 & \pi_1^*/3 & 0 & \cdots \\ \vdots & & & & \end{bmatrix}$$

Clearly each element on the right hand side of (7) is nonnegative and a look at the form of (7) will show that $\pi_i^*(n) > \pi_i^* > 0$.

A number of other results can be obtained based on equation (7) but they are left out for brevity.

3. Examples.

Example (i). Shows that the logarithmic distribution is i.d.

For a logarithmic distribution, $P_i = \theta^{i+1}/\{-(i+1) \log (1-\theta)\}, i = 0, 1, 2, \cdots$

where $0 < \theta < 1$. Hence, $P_i^* = P_i/P_0 = \theta^i/(i+1)$ and

$$\pi_i^* = i\theta^i/(i+1) - \sum_{j=1}^{i-1} (\theta^j/(j+1))\pi_{i-j}^* . \tag{8}$$

Change i to $i+1$ in (8) to get

$$\pi_{i+1}^* = (i+1)\theta^{i+1}/(i+2) - \sum_{j=1}^{i} \theta^j/(j+1)\pi_{i+1-j}^* . \tag{9}$$

Therefore

$$\pi_{i+1}^* - \theta\pi_i^* = \{(i+1)/(i+2) - i/(i+1)\}\theta^{i+1} + \sum_{j=1}^{i-1} \theta^j\pi_i^*(1/(j+1) - 1/(j+2)) - \theta/2\pi_i^*$$

or

$$\pi_{i+1}^* \geqq \theta/2\pi_i^*. \tag{10}$$

Since π_1^* is always nonnegative, equation (10) implies that $\pi_i^* \geqq 0$ for every i. Hence the logairthmic distribution is i.d.

Example (ii). The logarithmic distribution with zeros with $P_0 = a/(1+a)$ and $P_i = \theta^i/\{-i(1+a)\log(1-\theta)\}$, for $i = 1, 2, \cdots$ where $a > 0$ and $0 < \theta < 1$ is i.d. if and only if $a \geqq 1/\{-\log(1-\theta)\}$. To prove this, we write

$$P_i^* = P_i/P_0 = \theta^i/\{-ia\log(1-\theta)\}.$$

Formula (4) for $i = 2$ gives

$$\pi_1^* = 2P_2^* - (P_1^*)^2 \geqq 0. \tag{11}$$

On substituting for P_1^* and P_2^*, this inequality gives the condition given above Hence the condition is necessary. To prove its sufficiency, let us substitute the expression for P_i^* in formula (5). This gives the condition

$$\pi_i^* = i\theta^i/\{-ia\log(1-\theta)\} - \sum_{j=1}^{i-1} (\theta^j/\{-ja\log(1-\theta)\})\pi_{i-j}^* \geqq 0.$$

The fact that this inequality is satisfied can be verified through mathematical induction as in example (i) above. Hence, the condition $a \geqq 1/\{-\log(1-\theta)\}$ is necessary and sufficient for the logarithmic distribution with zeroes to be i.d.

Example (iii). Show that a distribution with probabilities

$$P_i = c\rho^{i+1}/(1-\rho^{i+1}),\ 0 < \rho < 1 \text{ is i.d.}$$

Here the condition $\pi_i^* \geqq 0$ reduces to showing that D_i given by the recursion formula $D_1 = 1/(1+\rho)$ and

$$D_i = i/(1+\rho+\cdots+\rho^i) - D_{i-1}/(1+\rho) - \cdots - D_1/(1+\rho+\cdots+\rho^{i-1})$$

are nonnegative. On changing i to $i+1$ and going through the arguments given above we get, $D_{i+1} \geqq \rho D_i/(1+\rho)$ and the proof of i.d. follows through induction.

Example iv. Show that the distribution $P_i = c\rho^{i+1}/\{(1-\rho)^i(1-\rho^{i+1})\}$, $0 < \rho < 1$ is i.d.

Proof is left to the reader. Distributions in examples (iii) and (iv) arise in queueing theory problems and were brought to the author's attention by Dr. Vincent Hodgson.