A NOTE ON CHARACTERIZING GEOMETRIC DISTRIBUTIONS

by

TaChen Liang
Wayne State University

Technical Report #90-13

Department of Statistics
Purdue University

March 1990

*

A NOTE ON CHARACTERIZING GEOMETRIC DISTRIBUTIONS *

by

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, MI 48202

Abstract

Let $X_1, \ldots, X_n$ be n mutually independent, nonnegative, integer–valued random variables such that for each $i = 1, \ldots, n, X_i$ has a probability function $f_i(x) = f(x; \theta_i)$, where $f(x; \theta_i) > 0$ for all $x = 0, 1, \ldots$. We show that conditional on the event that $X_i - X_{i+1} \geq 0$ for all $i = 1, \ldots, n-1$, the necessary and sufficient condition for the random variables $X_1 - X_2, \ldots, X_{n-1} - X_n$ and X_n to be mutually independent is that for each $i = 1, \ldots, n, X_i$ has a geometric distribution.

AMS 1980 Subject Classification: 62E10

Keywords: characterization, geometric distribution, conditional independence.

* This research was supported in part by the Office of Naval Research Contract N00014–88–K–0170 and NSF Grants DMS–8606964, DMS–8702620 at Purdue University.

1. Introduction

The problem of characterizing the geometric distributions have been intensively studied by many authors. Among the many papers in the literature, the reader is referred to Nagaraja (1988) for a nice review over this research field and to Nagaraja and Srivastava (1987), and Nagaraja, Sen and Srivastava (1989) for certain recent development on some related problems.

This paper is dealing with the problem of characterizing the geometric distributions via conditional independence. Let $X_1, \ldots, X_n$ denote n mutually independent, nonnegative, integer–valued random variables. It is assumed that for each $i = 1, \ldots, n, X_i$ has a probability function $f_i(x) = f(x; \theta_i)$, where $f(x; \theta_i) > 0$ for all $x = 0, 1, 2, \ldots$. That is, the n probability functions $f_i(x),\ i = 1, \ldots, n$, have the common form but their values depend on the respective parameter θ_i. Let E denote the event that $X_i - X_{i+1} \geq 0$ for all $i = 1, \ldots, n-1$. Given E, we seek a necessary and sufficient condition for $X_1 - X_2, \ldots, X_{n-1} - X_n$ and X_n to be conditionally mutually independent. We find that the necessary and sufficient condition is: for each $i = 1, \ldots, n, X_i$ has a geometric distribution. The main result is given in the next section.

2. Main Result

Let $X_1, \ldots, X_n$ denote n mutually independent, nonnegative integer–valued random variables. It is assumed that for each $i = 1, \ldots, n, X_i$ has a probability function $f_i(x)$, where $f_i(x) > 0$ for all $x = 0, 1, 2, \ldots$. Let E denote the event that $X_i \geq X_2 \geq \ldots \geq X_n$. A characterization of a geometric distribution via conditional independence is given as follows.

Lemma 1. Conditional on the event E, the following are equivalent.

(a) $X_1 - X_2$ and X_2 are independent.

(b) X_1 has a geometric distribution.

Proof: (Sufficiency) Suppose that $X_1 - X_2$ and X_2 are conditionally independent on the event E. For any nonnegative integers b and c, we then have the following:

$$P\{X_1 - X_2 \geq c,\ X_2 \geq b | E\} = P\{X_1 - X_2 \geq c | E\} P\{X_2 \geq b | E\}, \tag{1}$$

which is equivalent to

$$\begin{aligned}
&\frac{P\{X_1 - X_2 \geq 0,\ X_2 \geq b,\ X_2 - X_3 \geq 0, \ldots, X_{n-1} - X_n \geq 0\}}{P\{X_1 - X_2 \geq 0,\ X_2 - X_3 \geq 0, \ldots, X_{n-1} - X_n \geq 0\}} \\
&= \frac{P\{X_1 - X_2 \geq c,\ X_2 \geq b,\ X_2 - X_3 \geq 0, \ldots, X_{n-1} - X_n \geq 0\}}{P\{X_1 - X_2 \geq c,\ X_2 - X_3 \geq 0, \ldots, X_{n-1} - X_n \geq 0\}} \\
&= \frac{\sum_{x=b}^{\infty} P\{X_1 \geq x + c,\ X_3 \leq x,\ X_3 - X_4 \geq 0, \ldots, X_{n-1} - X_n \geq 0\} f_2(x)}{\sum_{x=0}^{\infty} P\{X_1 \geq x + c,\ X_3 \leq x,\ X_3 - X_4 \geq 0, \ldots, X_{n-1} - X_n \geq 0\} f_2(x)} \qquad (2) \\
&= \frac{\sum_{x=b}^{\infty} [1 - F_1(x + c - 1)] G(x) f_2(x)}{\sum_{x=0}^{\infty} [1 - F_1(x + c - 1)] G(x) f_2(x)} \equiv H(b, c),
\end{aligned}$$

where $G(x) = P\{X_3 \leq x,\ X_3 - X_4 \geq 0, \ldots, X_{n-1} - X_n \geq 0\}$, which is positive for all nonnegative integers x, since $f_i(y) > 0$ for all $y = 0, 1, 2, \ldots$, and for all $i = 1, \ldots, n$. The third equality in (2) is obtained by the assumption that $X_1, \ldots, X_n$ are mutually independent. Since the left–hand–side of (2) is independent of c, for each fixed b, $H(b, c)$ can be viewed as a constant function of the integer–valued variable c. Therefore, $H(b, c) - H(b, c + 1) = 0$, which yields

$$\begin{aligned}
0 = &\left\{ \sum_{x=b}^{\infty} [1 - F_1(x + c - 1)] G(x) f_2(x) \right\} \left\{ \sum_{x=0}^{\infty} [1 - F_1(x + c)] G(x) f_2(x) \right\} \\
&- \left\{ \sum_{x=b}^{\infty} [1 - F_1(x + c)] G(x) f_2(x) \right\} \left\{ \sum_{x=0}^{\infty} [1 - F_1(x + c - 1)] G(x) f_2(x) \right\} \\
&\equiv k(b, c)
\end{aligned}$$

for all nonnegative integers b and c. Thus,

$$\begin{aligned}
0 &= k(b + 1, c) - k(b, c) \\
&= [1 - F_1(b + c)] G(b) f_2(b) \left\{ \sum_{x=0}^{\infty} [1 - F_1(x + c - 1)] G(x) f_2(x) \right\} \\
&\quad - [1 - F_1(b + c - 1)] G(b) f_2(b) \left\{ \sum_{x=0}^{\infty} [1 - F_1(x + c)] G(x) f_2(x) \right\},
\end{aligned}$$

for all nonnegative integers b and c.

Letting $c = 0$ and from (4), after simple algebraic computation, we can obtain the following:

$$\frac{f_1(b)}{1 - F_1(b-1)} = 1 - B \text{ for all } b = 0, 1, 2, \ldots \tag{5}$$

where $B = \left\{ \sum_{x=0}^{\infty} [1 - F_1(x)] G(x) f_2(x) \right\} / \left\{ \sum_{x=0}^{\infty} [1 - F_1(x-1)] G(x) f_2(x) \right\}$. Note that B is a constant independent of the variable b and $0 < B < 1$. The assumption that $f_1(x) > 0$ for all $x = 0, 1, 2, \ldots$ and the fact of (5) together implies the following: X_1 has a geometric distribution with probability function $f_1(x)$ where $f_1(x) = B^x(1 - B)$, $x = 0, 1, 2, \ldots$

(Necessity). Assume that X_1 has a geometric distribution. Then the right–hand–side of (2) is a constant function of the variable c, which implies (1). Thus, conditional on the event E, $X_1 - X_2$ and X_2 are independent. □

We now assume that for each $i = 1, \ldots, n$, X_i follows a geometric distribution with probability function $f_i(x) = \theta_i^x(1 - \theta_i)$, $x = 0, 1, \ldots$, and $0 < \theta_i < 1$. Let $Y_i = X_i - X_{i+1}$, $i = 1, \ldots, n-1$, and $Y_n = X_n$. We have the following conditional independence property regarding the random variables $Y_1, \ldots, Y_n$.

Lemma 2 Conditional on the event $E, Y_1, \ldots, Y_n$ are mutually independent and for each $i = 1, \ldots, n$, Y_i has a geometric distribution with conditional probability function $f_i(y|E)$ where

$$f_i(y|E) = \left(\prod_{j=1}^{i} \theta_j \right)^y \left(1 - \prod_{j=1}^{i} \theta_j \right), \; y = 0, 1, 2, \ldots.$$

Proof: Straightforward computation yields that

$$\begin{aligned} P(E) &\equiv P(X_1 \geq X_2 \geq \ldots \geq X_n) \\ &= \prod_{i=2}^{n} [(1 - \theta_i)(1 - \theta_1 \ldots \theta_i)^{-1}]. \end{aligned} \tag{6}$$

Let $a_1, \ldots, a_n$ be n nonnegative integers. By the definition of the random variables

$Y_1, \dots, Y_n$,

$$P\{Y_i = a_i,\ i = 1, \dots, n\} = P\{X_i = \sum_{j=i}^{n} a_j,\ i = 1, \dots, n\} = \prod_{i=1}^{n} [(1-\theta_i)\theta_i^{\sum_{j=i}^{n} a_j}]. \tag{7}$$

From (6) and (7), it follows that

$$P\{Y_i = a_i, i = 1, \dots, n | E\} = P\{Y_i = a_i,\ i = 1, \dots, n\}/P\{E\} = \prod_{i=1}^{n} \left[\left(\prod_{j=1}^{i} \theta_j \right)^{a_i} \left(1 - \prod_{j=1}^{i} \theta_j \right) \right] \tag{8}$$

which implies that conditional on the event E, $Y_1, \dots, Y_n$ are mutually independent; also, Y_i has a conditional probability function $f_i(y|E) = \left(\prod_{j=1}^{i} \theta_j \right)^{y} \left(1 - \prod_{j=1}^{i} \theta_j \right)$, $y = 0, 1, 2, \dots$, for each $i = 1, \dots, n$. □

We now state our main result as a theorem as follows.

Theorem 1. Let $X_1, \dots, X_n$ be n mutually independent, nonnegative, integer–valued random variables such that X_i has a probability function $f_i(x) = f(x; \theta_i)$ where $f(x; \theta_i) > 0$ for all $x = 0, 1, 2, \dots$. Let $E = \{X_1 \geq X_2 \geq \dots \geq X_n\}$. Then, the following are equivalent.

(a) Conditional on the event E, $X_1 - X_2, \dots, X_{n-1} - X_n$ and X_n are mutually independent.

(b) For each $i = 1, \dots, n$, X_i has a geometric distribution with probability function $f_i(x) = \theta_i^x (1 - \theta_i)$, $x = 0, 1, 2, \dots$ for some θ_i, $0 < \theta_i < 1$.

Proof: (Sufficiency) Conditional on the event E, the mutual independence among the random variables $X_1 - X_2, \dots, X_{n-1} - X_n$ and X_n implies the conditional independence between $X_1 - X_2$ and X_2 on the event E. Then by Lemma 1, X_1 follows a geometric distribution with $\theta_1 = B$ where B is defined in the proof of Lemma 1. By the assumption, the probability functions of $X_1, \dots, X_n$ have the common form, which implies that X_i also has a geometric distribution for each $i = 2, \dots, n$.

(Necessity). See Lemma 2.

References

Nagaraja, H.N. (1988). Record values and related statistics – a review. *Commun. Statist.-Theory Meth.*, **17(7)**, 2223–2238.

Nagaraja, H.N., Sen, P. and Srivastava, R.C. (1989). Some characterizations of geometric type distributions based on record values. *Statistische Hefte*, **30**, 147–155.

Nagaraja, H.N. and Srivastava, R.C. (1987). Some characterizations of geometric type distributions based on order statistics. *J. Statist. Plann. Inference*, **17**, 181–191.