PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 10, Fasc. 1 (1989), p. 143–147

CHARACTERISATIONS OF THE GEOMETRIC DISTRIBUTION USING DISTRIBUTIONAL PROPERTIES OF THE ORDER STATISTICS

BY

SITADRI NATH BAGCHI (RENO, NV)

Abstract. Let $X_1, X_2, \ldots, X_N$ be independent, identically distributed, non-negative integer valued random variables and let

$$X_{1:N} \leqslant X_{2:N} \leqslant \ldots \leqslant X_{N:N}$$

denote the corresponding order statistics. If $\mathrm{P}(X_{1:N} \geqslant k) = \mathrm{P}(X_1 \geqslant Nk)$ for all $N \geqslant 1$ and $k = 1$, then the distribution of X_1 is geometric. We modify this distributional property of the order statistics in several directions to obtain characterisations of the geometric distribution. We provide examples to show that the assumptions, in some respect, cannot be weakened any further.

1. Introduction. Let $X_1, X_2, \ldots, X_n$ be independent, identically distributed, non-negative integer valued random variables and let $X_{1:n} \leqslant X_{2:n} \leqslant \ldots \leqslant X_{n:n}$ be the coresponding order statistics. Much has been done in the way of characterising the geometric distribution by the independence of functions of the order statistics. In our work we propose some characterisations using distributional properties of the order statistics.

Consider the condition

$$\mathrm{P}(X_{1:N} \geqslant k) = \mathrm{P}(X_1 \geqslant Nk). \tag{1}$$

It is easy to show (sse Galambos [1]) that if (1) holds for all $N \geqslant 1$ and $k = 1$, then X_1 has a geometric distribution, i.e.,

$$\mathrm{P}(X_1 = k) = p(1-p)^k, \tag{2}$$

for some $0 \leqslant p \leqslant 1$ and $k = 0, 1, 2, \ldots$ Gupta [2] showed that if the common distribution function $F(x)$ of $X_1, X_2, \ldots, X_n$ is continuous and satisfies some differentiability condition at $x = 0$ and if

$$\mathrm{P}(X_{1:n} \geqslant x) = \mathrm{P}(X_1 \geqslant nx) \tag{3}$$

for all $x > 0$ and for some fixed but arbitrary n, then the distribution is exponential. Here we obtain characterisations of the geometric distribution based on (1). We say that two natural numbers m and n are *incommensurable* if one is not a rational power of the other, i.e., if $\log_m n$ is irrational.

2. Main results.

THEOREM 2.1. *Let $X_1, X_2, \ldots, X_N$ be i.i.d non-negative integer valued random variables such that* $P(X_1 = 0) < 1$. *If* (1) *holds for all k and for two incommensurable values $1 < m < n$ of N, then X_1 has a distribution given by formula* (2).

Proof. The hypothesis (1) can also be written as

$$[P(X_1 \geqslant k)]^N = P(X_1 \geqslant Nk). \tag{4}$$

We claim that

$$[P(X_1 \geqslant k)]^{N^i} = P(X_1 \geqslant N^i k) \quad \text{for all } i, k \geqslant 1 \text{ and for } N = m, n. \tag{5}$$

Indeed, if $i = 1$, then (5) reduces to (4). Assume that (5) is true for $i = 1, 2, \ldots, j$. Then,

$$P(X_1 \geqslant N^{j+1} k) = P(X_1 \geqslant N^j Nk) = [P(X_1 \geqslant Nk)]^{N^j} = [P(X_1 \geqslant k)]^{N^{j+1}}.$$

Finally, for $i, j \geqslant 1$, we have

$$P(X_1 \geqslant m^i n^j k) = [P(X_1 \geqslant m^i k)]^{n^j} = [P(X_1 \geqslant k)]^{m^i n^j}. \tag{6}$$

Let $G(k) = P(X_1 \geqslant k)$ and let $G(1) = q$. Since $G(m^i) = [G(1)]^{m^i}$ for all i and since $\lim G(k) = 0$, as $k \to \infty$ it follows that $q < 1$. We shall use the following lemma, the proof of which is given in the Appendix.

LEMMA 2.3. *Given two incommensurable natural numbers m and n, a natural number k not of the form $m^i n^j$ and $\varepsilon > 0$,*

(i) *there exist integers i and j, not both positive, such that*

$$\ln k < i \ln m + j \ln n < \ln(k+\varepsilon) \tag{7}$$

and

(ii) *there exist integers i' and j', not both positive, such that*

$$\ln(k-\varepsilon) < i' \ln m + j' \ln n < \ln k. \tag{8}$$

Remark. If such that integers i and j exist, both of them cannot be negative.

We shall prove that $G(k) = q^k$ for all $k \geqslant 1$. Suppose that, for some k, $G(k) = q^{k+\varepsilon}$ for some $\varepsilon \neq 0$. From (6) we know that k is not of the form $m^i n^j$. Suppose first that $\varepsilon > 0$. By (7) of Lemma 2.3, there exist i and j, not both positive, such that $\ln k < i \ln m + j \ln n < \ln(k+\varepsilon)$. If $i \leqslant 0$ and $j > 0$, then

$$km^{-i} < n^j < (k+\varepsilon) m^{-i}. \tag{9}$$

Hence $G(km^{-i}) \geqslant G(n^j)$. But $G(n^j) = q^{n^j} > q^{(k+\varepsilon)m^{-i}} = G(km^{-i})$, which is a contradiction. We can obtain a similar contradiction if $i > 0$ and $j \leqslant 0$. Finally, if $\varepsilon < 0$, then we use (8) of Lemma 2.3 to arrive at a similar contradiction. Hence $G(k) = q^k$ for all $k \geqslant 1$ and $P(X_1 = k) = G(k) - G(k+1) = p(1-p)^k$. Hence X_1 has a distribution given by (2).

Example 2.4. Can we weaken the hypothesis of Theorem 2.1 by assuming that (1) holds for all values of $k \geqslant 1$ and for a fixed value of N? If such a result were true, then it would be an appropriate extension of Gupta's [2] result to the discrete case. However, the following example shows that that is not possible. Here the distribution of X_1 is not geometric, but $X_1, \ldots, X_N$ satisfies (1) for all $k \geqslant 1$ and $N = 2$. We specify the distribution by specifing $G(k)$ for $k \geqslant 1$, and we define $G(k)$ inductively. Let $G(0) = 1$ and $G(1) = q$ for some q such that $0 < q < 1$. Let $G(2) = q^2$ and $G(3) = q^{2+\delta}$, where $0 < \delta < 1$. Suppose that $G(0), G(1), \ldots, G(k)$ have been defined. If $k+1 = 2n$, then define $G(k+1) = G^2(n)$. If $k = 2m$, then $m+1 < k$. Suppose that $G(m) = q^\alpha$ and $G(m+1) = q^\beta$. Then define $G(k+1) = q^{\alpha+\beta}$. It can be checked easily that $G(k)$ is a strictly decreasing sequence. The fact that, for $m \to \infty$, $\lim G(2^m) = \lim q^{2^m} = 0$ implies that $\lim G(m) = 0$. Thus $P(X_1 = k) = G(k) - G(k+1)$ defines a valid probability distribution. From the definition of $G(k)$ it follows that $G(2k) = G^2(k)$ for all $k \geqslant 1$, i.e., this distribution satisfies (1) for $N = 2$ and for all k. However, X_1 does not have a geometric distribution.

Remark. One can construct distribution along the lines of the above example which satisfies (1) for any arbitrary but fixed N. We chose $N = 2$ for the simplicity of notation.

It is also natural to ask whether the underlying distribution could be geometric if (1) holds for all N, but for some value of $k > 1$. The following simple example shows that such a result is not true.

Example 2.5. Let $0 < k < 1$ and let $P(X_1 = k) = p(1-p)^k$ for $k \geqslant 2$ and $P(X_1 = 0)$, $P(X_1 = 1)$ be arbitrary but such that $P(X_1 = 0) + P(X_1 = 1) = 1 - q^2$, where $q = 1-p$. Then $G(k) = q^k$ for all $k \geqslant 2$ and (1) is satisfied for all N and all $k \geqslant 2$. Clearly the distribution need not be geometric.

However, the following result is true.

PROPOSITION 2.6. *Let m and l be two co-prime numbers. Suppose that* (1) *is satisfied for all $N \geqslant 1$ and $k = m$. Suppose also that X_1 has "lack of memory at age l", i.e., $P(X_1 \geqslant l) > 0$ and*

$$P(X_1 = l+i \mid X_1 \geqslant l) = P(X_1 = i) \quad \text{for } i \geqslant 0. \tag{10}$$

Then X_1 has a geometric distribution.

Proof. One can write (10) as

(11) $$P(X_1 \geqslant l+i) = P(X_1 \geqslant i) P(X_1 \geqslant l).$$

Thus, in particular, $G(2l) = G^2(l)$. Applying induction on n, it can be shown that $G(nl) = G^n(l)$ and that

(12) $$G(nl+i) = G(i) G^n(l) \quad \text{for all } i, n \geqslant 0.$$

Since m and l are co-prime, there exist positive integers α and β such that $\alpha m - \beta l = 1$. Hence,

(13) $$n\alpha m = n\beta l + n \quad \text{for all } n \geqslant 0.$$

Hence $G(n\alpha m) = G(n\beta l + n)$. Due to our hypothesis $G(n\alpha m) = [G(m)]^{n\alpha}$ and, due to (12), $G(n\beta l+n) = G^{n\beta}(l) G(n)$. Therefore $G^{n\alpha}(m) = G^{n\beta}(l)\ G(n)$ for all $n \geqslant 0$ and

(14) $$G(n) = q^n, \quad \text{where } q = G^{\alpha}(m) G^{-\beta}(l).$$

Since $\alpha m > \beta l$, it follows that $G^{\alpha}(m) = G(m\alpha) \leqslant G(\beta l) = G^{\beta}(l)$. Hence $q \leqslant 1$. But since $G(n) = P(X_1 \geqslant n) \to 0$ as $n \to \infty$, it follows that $q < 1$. But (14) is a characterising property of the geometric distribution.

3. Appendix.

Proof of Lemma 2.3. We need the following result to prove this lemma:

RESULT 3.1. *Let A be a closed additive subgroup of $\boldsymbol{R}$. Then either $A = \boldsymbol{R}$ or A is generated by some number θ, i.e., $A = \{n\theta\colon n\varepsilon\}$.*

We provide the proof of (i) only since the proof of (ii) is quite similar. Consider the additive subgroup of A described by

$$A = \{i \ln m + j \ln n\colon i, j\varepsilon\}.$$

Since $\bar{A}$ is a closed subgroup of $\boldsymbol{R}$ either $\bar{A} = \boldsymbol{R}$ or $\bar{A} = \{n\theta\colon n\varepsilon\}$ for some real number θ. If the latter is the case, then $\ln m = m_1 \theta$ and $\ln n = n_1 \theta$ for some integers m_1 and n_1. But then $\ln m / \ln n = m_1/n_1$, which is a rational number contradicting the hypothesis that m and n are incommensurable. Hence $\bar{A} = \boldsymbol{R}$, i.e., A is dense in $\boldsymbol{R}$. Hence, for some integers i and j, $i \ln m + j \ln n$ is in the interval $(\ln k, \ln(k+\varepsilon))$.

Clearly, both i and j cannot be negative. We now show that it is possible to choose i and j such that both are not positive. Suppose that that is not the case. Let (ε_r) be a sequence of positive numbers such that $\varepsilon_r \leqslant \varepsilon$ for each $r \geqslant 1$ and $\varepsilon_r \to 0$ as $r \to \infty$. It follows from the last section that, for every $r \geqslant 1$, there exists a pair i_r, j_r such that $\ln k < i_r \ln m + j_r \ln n < \ln(k+\varepsilon_r)$. Since $\varepsilon_r \to 0$ as $r \to \infty$, it follows that

(15) $$\lim_{r\to\infty} m^{i_r} n^{j_r} = k.$$

But, due to our assumption, $i_r, j_r > 0$ for every $r \geqslant 1$. Therefore, the only way (15) can be true is if $k = m^{i_r} n^{j_r}$ for sufficiently large r. This contradicts the assumption that k is not of the form $m^i n^j$. This completes the proof of the lemma.

REFERENCES

[1] J. Galambos, *Characterizations of probability distributions by properties of order statistics, II*, Statistical Distributions in Scientific Work, ed. G. P. Patilet al., Riedl, Dordrecht 1975.

[2] R. C. Gupta, *A characteristic property of the exponential distribution*, Saṅkhyā, B, 35 (1973), p. 365–366.

Department of Mathematics
Univeristy of Nevada Reno
Reno, Nevada 89502
U.S.A.

Received on 29. 4. 1987