Sankhyā : The Indian Journal of Statistics
1984, Volume 46, Series B, Pt. 1, pp. 108–109.

NOTES

ON CHARACTERIZING DISTRIBUTIONS FOR WHICH THE SECOND RECORD VALUE HAS A LINEAR REGRESSION ON THE FIRST

By R. M. KORWAR
University of Massachusetts
and
Indian Statistical Institute

SUMMARY. Let $\{X_n\}$ be a sequence of independent and identically distributed random variables taking non-negative integer values and having finite expectation, and let R_1 and R_2 be the first two record values. In this note we show that if R_2 has a linear regression on R_1, then the distribution of X_1 must be one of three kinds, including the geometric (and conversely) : this extends the work of Srivastava (1979) where $E(R_2-R_1\,|\,R_1)$ = constant, almost surely, is shown to characterize the two-parameter geometric.

Let $\{X_n\}$ be a sequence of independent and identically distributed random variables taking non-negative integer values and having finite expectation, with $P[X_1 = j] = p_j$ for $j = 0, 1, \ldots, m$. In order that R_2 below be defined we assume that m is a positive integer or possibly ∞. Let $N(1) = 1$ and $N(2) = \min\{j : X_j > X_1\}$; $R_1 = X_{N(1)} = X_1$ and $R_2 = X_{N(2)}$. R_1 and R_2 are called the first two record values of $\{X_n\}$. Srivastava (1979) computed that

$$E(R_2-R_1\,|\,R_1 = i) = \sum_{j=1}^{m-i} jp_{i+j} \Big/ \sum_{j=i+1}^{m} p_j \qquad \ldots \quad (1)$$

for $i = 0, \ldots, m-1$; in case $m = \infty$, the preceding and similar statements below are to be interpreted in an obvious manner. He also showed that if R_2-R_1 has constant regression on R_1, then X_1 has a geometric distribution (and conversely). If now we assume that

$$E(R_2-R_1\,|\,R_1 = i) = \alpha+\beta i \text{ a.s., } i < m \qquad \ldots \quad (2)$$

where α and β are constants, then we easily obtain, on first equating the right hand sides of (1) and (2) for the value $i+1$, clearing fractions and subtracting the relation thus obtained from the same relation for the value i, that

$$(1+\beta) \sum_{j=i+1}^{m} p_j = \{\alpha+(i+1)\beta\}p_{i+1}, \quad i = 0, \ldots, m-2. \qquad \ldots \quad (3)$$

Subtracting (3) for the value $i+1$ from (3), we obtain

$$p_{i+2}/p_{i+1} = (\alpha+i\beta-1)/\{\alpha+(i+2)\beta\} \qquad \ldots \quad (4)$$

AMS subject classification : Primary 62E10 ; Secondary 60E05.

Key words and phrases : Characterization, two-parameter geometric distribution, generalized hypergeometric distribution, linear regression, record values.

for $i = 0, \ldots, m-2$. Setting $i = 0$ in (3), we get

$$p_1 = (1-p_0)(1+\beta)/(\alpha+\beta) \quad \ldots \quad (5a)$$

and (4) implies that

$$p_j = [p_1(\alpha-1) \ldots (\alpha+(j-2)\beta-1)]/[(\alpha+\beta) \ldots (\alpha+j\beta)] \quad \ldots \quad (5b)$$

for $j = 2, \ldots, m$. Now, since $R_2 - R_1 \geqslant 1$, (2) implies that $\alpha+i\beta \geqslant 1$ for $i < m$, and in particular $\alpha \geqslant 1$. Relation (5a), inequality $\alpha+\beta \geqslant 1$ and the fact $m \geqslant 1$ show that $\beta > -1$. So we need only consider the three cases $-1 < \beta < 0$, $\beta = 0$ and $\beta > 0$. If $-1 < \beta < 0$, then m is necessarily finite and, in such a case, (3) implies (take $i = m-1$) that $\alpha+(m-1)\beta = 1$ and we can rewrite (5b) as

$$p_j = c(1-p_0)(m-1)^{(j-1)}/(c+m-1)^{(j)}, \quad j = 1, \ldots, m \quad \ldots \quad (6)$$

where $c = -1-1/\beta$, and $x^{(j)} = x(x-1) \ldots (x-j+1)$, $x^{(0)} = 1$. Conversely, if m is finite, then $\alpha+(m-1)\beta = 1$; in such a case, $\beta = 0$ leads to a distribution concentrated on $\{0, 1\}$, as (5a) indicates in view of $\alpha = 1$ then. The case $\beta = 0$ and $m = \infty$ corresponds to the two-parameter geometric distribution for X_1, as indicated by (5a) and (5b); precisely,

$$p_j = (1-p_0)(\alpha-1)^{j-1}/\alpha^j, \quad j = 1, 2, \ldots, \text{ad inf.}$$

Finally, if $\beta > 0$, then m is necessarily ∞ and the p_j are given by (5a) and (5b) for $j = 1, 2, \ldots$ ad inf. We may consider (6) as a Polya-Eggenberger type of distribution and (5) as a generalized hypergeometric (cf. Kemp and Kemp (1956)).

We can easily verify that for each distribution of the above three types, (2) holds.

Acknowledgements. The author wishes to sincerely thank the referee for constructive suggestions in shortening the paper. This research was supported by a grant from the Air Force Office of Scientific Research, AFSC, USAF.

References

Kemp, C. D. and Kemp, A. W. (1956): Generalized hypergeometric distributions. *J. Roy. Stat. Soc.*, Ser B, **18**, 202-11.

Srivastava, R. C. (1979): Two characterizations of the geometric distribution by record values. *Sankhyā*, Ser. B., **40**, 276-78.

Paper received : August, 1981.
Revised : April, 1983.