A CHARACTERIZATION OF THE GEOMETRIC DISTRIBUTION

T. S. FERGUSON, University of California, Los Angeles

1. Introduction. Certain distributions may be characterized by conditions asserting the independence of suitable statistics. The most renowned of these characterization theorems is that first proved by Kac [5] which states: *if random variables X and Y are independent, then $X+Y$ and $X-Y$ are independent if and only if, both X and Y have normal distributions with a common variance.*

A similar characterization of the gamma distribution has been proved by Hogg [4] and Lukacs [6]. It is the purpose of this paper to investigate the extent to which the geometric distribution may be characterized by analogous assumptions of independence.

The starting point for this investigation was the observation in [1] that if X and Y are independent random variables each following the same geometric law with probability mass function

$$f(k) = (1 - p)p^k \qquad k = 0, 1, 2, \cdots, \tag{1}$$

where $0<p<1$, then the random variables $\min(X, Y)$ and $X-Y$ are stochastically independent. It was also observed, however, that if X and Y are independent random variables having the same exponential distribution with density

$$\begin{aligned} f(z) &= \lambda e^{-\lambda(z-\theta)} && \text{if} \quad z > \theta \\ &= 0 && \text{if} \quad z < \theta, \end{aligned} \tag{2}$$

where $\lambda>0$, then again the random variables $\min(X, Y)$ and $X-Y$ are stochastically independent. Thus, the independence of $\min(X, Y)$ and $X-Y$ for independent X and Y will not by itself characterize either the geometric or exponential distributions, so that a complete analogue of the result of Kac cannot be valid. Therefore, in order to obtain a characterization of either the geometric or the exponential distribution, it is necessary to restrict the distributions in some way. In [2], the following result was obtained: *if X and Y are independent random variables with absolutely continuous distributions, then* $\min(X, Y)$ *and* $X-Y$ *are independent if, and only if, both X and Y have exponential distributions of the form* (2) *with a common value of θ (but possibly different values of λ).*

For a formulation of the problem which contains the normal, the gamma, the exponential, the geometric, and other distributions in a single statement, see [1] or [2]. In this paper the distributions of X and Y are restricted to be discrete.

The term *geometric distribution* signifies in this paper the distribution of any random variable X for which there is a real number θ and a positive real number c such that the random variable $(X-\theta)/c$ has the distribution of formula (1). The parameters θ and c are respectively location and scale parameters of the distribution of X. The parameter p will be referred to as the geometric parameter. A random variable X is said to be degenerate at x_0 if $P(X=x_0)=1$. In the next section, it is shown that if X and Y are independent, nondegenerate, discrete random variables, then $\min(X, Y)$ and $X-Y$ are independent if, and only if, X and Y both have geometric distributions with common location and scale parameters, but with possibly different geometric parameters.

With such a theorem, the study of the independence of $\min(X, Y)$ and $X-Y$ for independent, discrete X and Y is easily completed with the aid of the following simple theorem which takes care of the degenerate cases.

THEOREM 1. *If X is degenerate at x_0, then* $\min(X, Y)$ *and* $X-Y$ *are independent if, and only if, either* (i) *Y is degenerate or* (ii) $P(Y\geqq x_0)=1$.

Proof. The "if" part of the theorem is immediate, since degenerate random variables are independent of any random variables, and since both (i) and (ii) imply that $\min(X, Y)$ is degenerate. To prove the "only if" part, suppose that Y is nondegenerate and that $P(Y\geqq x_0)<1$. Then, there exists a number $m<x_0$ such that $0<P(Y<m)<1$. Using the independence of $\min(x_0, Y)$ and x_0-Y, we have

$$\begin{aligned} P(Y<m) &= P(\min(x_0, Y)<m) = P(\min(x_0, Y)<m \mid x_0-Y>x_0-m) \\ &= P(Y<m \mid Y<m) = 1. \end{aligned}$$

This contradiction completes the proof.

The analogous theorem where Y is degenerate is also valid, since if $\min(X, Y)$ is independent of $X-Y$, it is automatically independent of $Y-X$.

2. The characterization theorem. We first present two lemmas. The first lemma may be found as a special case of a lemma of Ghurye [3, Lemma 1] proved by means of Fourier transforms. For the sake of completeness and simplicity we give a proof of Lemma 1 involving only elementary notions.

LEMMA 1. *If X and Y are independent, and if X and $X-Y$ are independent, then X is degenerate.*

Proof. Suppose that X is not degenerate. Then, there exist four points $x_1<x_2<x_3<x_4$ such that $P\{x_1\leqq X\leqq x_2\}>0$ and $P\{x_3\leqq X\leqq x_4\}>0$. For any other variable Y and all numbers y

$$P\{Y - X \leqq y, x_1 \leqq X \leqq x_2\} \leqq P\{Y \leqq y + x_2, x_1 \leqq X \leqq x_2\} \tag{3}$$

and

$$P\{Y - X \leqq y, x_3 \leqq X \leqq x_4\} \geqq P\{Y \leqq y + x_3, x_3 \leqq X \leqq x_4\}. \tag{4}$$

Inequality (3) and the independence assumptions imply

$$P\{Y - X \leqq y\} \leqq P\{Y \leqq y + x_2\} \tag{5}$$

for all y. Similarly, from inequality (4),

$$P\{Y \leqq y + x_3\} \leqq P\{Y - X \leqq y\} \tag{6}$$

for all y, which in conjunction with (5) implies that Y has no mass anywhere. This contradiction completes the proof.

LEMMA 2. *Let X and Y be independent nondegenerate random variables and suppose that $U=\min(X, Y)$ and $V=X-Y$ are independent. If $P(X=x)>0$ and $P(Y>x)>0$, then $P(Y=x)>0$. Similarly, if $P(Y=y)>0$ and $P(X>y)>0$, then $P(X=y)>0$.*

Proof. Suppose $P(Y=x)=0$. Then if $P(X=x)>0$ and $P(Y>x)>0$,

$$P(V \leqq 0) = P(V \leqq 0 \mid U = x) = P(X \leqq Y \mid X = x, Y \geqq x) = 1. \tag{7}$$

Thus, $P(X\leqq Y)=1$, so that U is equal to X with probability 1. Since now X is independent of $X-Y$, Lemma 1 will provide a contradiction, proving that $P(Y=x)>0$. The last statement follows by symmetry since if U is independent of V it is also independent of $-V$.

A number x is said to be a possible value of a discrete random variable X, if $P(X=x)>0$. Using this terminology, the conclusions of Lemma 2 may be stated for discrete variables as: *Every possible value of X less than some possible value of Y is a possible value of Y, and every possible value of Y less than some possible value of X is a possible value of X.*

THEOREM 2. *Let X and Y be independent, nondegenerate, discrete random variables. Then, $U=\min(X, Y)$ and $V=X-Y$ are independent if, and only if, both X and Y have geometric distributions with the same location and scale parameters.*

Proof. First, suppose that U and V are independent; then for all u and all $v\geqq 0$,

$$\begin{aligned} P(U = u)P(V = v) &= P(\min (X, Y) = u, X - Y = v) \\ &= P(X = u + v)P(Y = u). \end{aligned} \tag{8}$$

Lemma 2 implies that $P(V=0)\neq 0$, so that by putting $v=0$ in equation (8) we may solve for $P(U=u)$. Substituting this value into equation (8) gives for all u and all $v \geqq 0$

$$P(X = u)P(Y = u)P(V = v) = P(X = u + v)P(Y = u)P(V = 0). \tag{9}$$

Lemma 2 and nondegeneracy imply that there are at least two distinct numbers, $x_0 < x_1$, which are possible values simultaneously for X and Y. Two equations may be obtained by letting u assume the values x_0 and x_1 in (9), from which $P(V=v)$ may be eliminated, yielding for all $v > 0$,

$$P(X = x_1 + v) = \left[\frac{P(X = x_1)}{P(X = x_0)}\right] P(X = x_0 + v). \tag{10}$$

This implies, by induction, that

$$P(X = n(x_1 - x_0) + x_0) = \left[\frac{P(X = x_1)}{P(X = x_0)}\right]^n P(X = x_0). \tag{11}$$

Since the total mass of the distribution of X must be finite, $P(X=x_0) > P(X=x_1)$. But since x_0 and x_1 were arbitrary numbers which were possible values of both X and Y, we see that there can be at most a finite number of possible values of X which are less than x_1. We now suppose that x_0 is the smallest of the possible values of X and that x_1 is the next smallest. Then equation (10) inductively implies that the only possible values of X are $n(x_1 - x_0) + x_0$, $n = 0, 1, 2, \cdots$ and equation (11) implies that X has a geometric distribution on these values. By symmetry, Y also has a geometric distribution, and Lemma 2 implies that Y must have the same set of possible values, and hence the same location and scale parameters.

Now, suppose that X and Y have geometric distributions with the same location and scale parameters, and geometric parameters p_1 and p_2 respectively. A change of location and scale will not affect the independence or nonindependence of U and V, so we may choose the variables to be defined on the nonnegative integers as in equation (1). Then, for all u and all $v \geqq 0$,

$$P\{U = u, V = v\} = P\{X = u + v\}P\{Y = u\} = (1 - p_1)(1 - p_2)p_1^{u+v}p_2^{u},$$

while for all u and $v < 0$,

$$P\{U = u, V = v\} = P\{X = u\}P\{Y = u - v\} = (1 - p_1)(1 - p_2)p_1^{u}p_2^{u-v}.$$

Thus $P\{U=u, V=v\}$ obviously factors into a product of a function of u, say $P\{U = u\} = (1 - p_1p_2)(p_1p_2)^u$, and a function of v, say $P\{V = v\} = (1-p_1)(1-p_2)(1-p_1p_2)^{-1}p_1^v$ for $v \geqq 0$, and

$$P\{V = v\} = (1 - p_1)(1 - p_2)(1 - p_1p_2)^{-1}p_2^{-v} \quad \text{for} \quad v < 0,$$

proving the independence of U and V.

References

1. T. S. Ferguson, Location and scale parameters in exponential families of distributions, Ann. Math. Statist., 33 (1962) 986–1001.

2. ———, A characterization of the exponential distribution, Ann. Math. Statist., 35 (1964) 1199–1207.

3. S. G. Ghurye, Random functions satisfying certain linear relations, Ann. Math. Statist., 25 (1954) 543–554.

4. R. V. Hogg, On ratios of certain algebraic forms, Ann. Math. Statist., 22 (1951) 567–572.

5. M. Kac, On a characterization of the normal distribution, Amer. J. Math., 6 (1939) 726–728.

6. E. Lukacs, A characterization of the gamma distribution, Ann. Math. Statist., 26 (1955) 319–324.