CHARACTERIZATION OF GEOMETRIC AND EXPONENTIAL DISTRIBUTIONS

By Gordon B. Crawford

Boeing Scientific Research Laboratories, Seattle, Washington

1. Introduction. Consider the following property of two independent random variables X and Y:

$$W = \min(X, Y) \quad \text{independent of} \quad X - Y.$$

In [2] Thomas S. Ferguson proves that if X or Y have a discrete part, then this property implies that for suitable constants a, b, $b(X - a)$ and $b(Y - a)$ have (possibly different) geometric distributions; i.e.,:

$$\begin{aligned} P[b(X - a) = n] &= (1 - p)p^n, \qquad n \geqq 0, \\ &= 0 \qquad\qquad\quad \text{otherwise.} \end{aligned}$$

In [3], by the same author, it is shown that if X and Y are absolutely continuous, then for suitable a, $(X - a)$ and $(Y - a)$ have possibly different exponential distributions, i.e.:

$$\begin{aligned} P(X - a \geqq c) &= e^{-c/\lambda}, \qquad c \geqq 0, \\ &= 1 \qquad\qquad \text{otherwise.} \end{aligned}$$

In [1] A. P. Basu gets the same results as [3] under slightly different conditions. It is assumed that X and Y are identically distributed with absolutely continuous distribution $F(\cdot)$; $F(0) = 0$; and the seemingly weaker independence condition:

W, the first order statistic, is independent of the difference $|X - Y|$ of the order statistics.

Basu's result may be obtained from a paper [4] by G. S. Rogers by taking the logarithms of the random variables considered in [4]. Rogers' paper is interesting in that the proof requires only that the regression of $e^{-|X-Y|}$ on W is constant.

In the concluding remarks of [3] Ferguson points out the unsettled problem that arises if X or Y has a singular part. We intend to resolve this problem, assuming that W is independent of $X - Y$.

The main result here is that if the independent random variables X and Y have the property that W is independent of $X - Y$, then X and Y are both geometric random variables or they are both exponential random variables.

We attempt to avoid some measure-theoretic difficulties by working with the distribution functions instead of the densities. The method is different from those mentioned above; all of the results of Ferguson are achieved at little additional expense. Lemmas 1 and 2 are consequences of the asserted independence of W and $X - Y$. Theorem 1 gives a condition which is equivalent to discreteness,

Received 9 August 1965; revised 25 May 1966.

and Theorem 2 shows that if this condition is not satisfied, then $X - a$ and $Y - a$ are exponential. Theorem 3 gives Ferguson's result of [2]; if X and Y are discrete, then $b(X - a)$ and $b(Y - a)$ are geometric. Throughout we assume that X and Y are non-degenerate.

2. Theorems and proofs, condition C. Hereafter we assume that X and Y are non-degenerate independent random variables with underlying probability measure P, and with the property that $X - Y$ is independent of the order statistic $W = \min(X, Y)$.

In the sequel these conditions will be abbreviated by saying that X and Y satisfy condition C.

Note that if X and Y satisfy condition C, then $|X - Y|$ and W are independent.

Let $P_W(\cdot)$ denote the probability measure on the real line induced by W.

LEMMA 1. *Under condition* C:

$$P(X > Y) > 0, \qquad P(Y > X) > 0, \qquad P(W > 0) > 0.$$

PROOF. Suppose $P(Y > X) = 0$. Then W is a.s. equal to X, in this case X and $-Y$ are independent and X is independent of $X - Y$, hence the characteristic functions satisfy

$$\Phi_{-Y}(t) = \Phi_{(X-Y)}(t)\Phi_{-X}(t) = \Phi_{-Y}(t)\Phi_X(t)\Phi_{-X}(t).$$

We know that $|\Phi_X(t)| \leqq 1$, it follows from the above that in a neighborhood of the origin $|\Phi_X(t)| = 1$, hence X is degenerate, which is a contradiction to condition C.

Similarly we may show $P(X > Y) > 0$.

Assume $P(W > 0) = 0$.

Then at least one of the random variables X and Y is a.s. non-positive. We will assume that the right hand end point of the range of X is non-positive and is no greater than the right hand end point of the range of Y.

Let $\delta > 0$ be such that $P(X - Y > \delta) > 0$. Let $-\eta$, $(\eta > 0)$, be the right hand end point of the range of X (and consequently the right hand end point of the range of W). Then for $0 < \epsilon < \delta$

$$P(-\eta \geqq W > -\eta - \epsilon) > 0,$$

$$P(X - Y > \epsilon) > 0,$$

but $\{(x, y): -\eta \geqq \min(x, y) > -\eta - \epsilon, x - y > \epsilon\} = \varnothing$, hence W and $X - Y$ cannot be independent. ∎

If X and Y are independent and $X - Y$ is independent of W, then the same is true for the random variables $\tilde{X} = (X - a)$ and $\tilde{Y} = (Y - a)$. Thus we have

COROLLARY TO LEMMA 1. *Under condition* C *the ranges of the random variables* X, Y, *and* W *are unbounded to the right.*

LEMMA 2. *If* X *and* Y *satisfy condition* C, *then*

$$P(X > Y)[P(X > \theta + c)/P(X > \theta)] + P(Y > X)[P(Y > \theta + c)/P(Y > \theta)] = P(|X - Y| > c)$$

for all $c \geqq 0$ *except on a* θ *set of* P_W *measure* 0. (*The exceptional set not depending on* c.)

Proof. First we prove the assertion for fixed c. The assertion follows immediately on a countably dense set, and hence for all c, since the probabilities involved, when considered as functions of c, are right continuous. Now, for fixed $c \geqq 0$:

$$P(|X - Y| > c \mid W = \theta) = h(c) = P(|X - Y| > c)$$

where $h(c)$ is independent of θ almost surely $P_W(\theta)$.

$$\begin{aligned} h(c) &=_{\text{a.s.}} P(X - Y > c \mid W = \theta, X > Y)P(X > Y \mid W = \theta) \\ &\quad + P(Y - X > c \mid W = \theta, Y > X)P(Y > X \mid W = \theta) \\ &\quad + P(Y - X > c \mid W = \theta, Y = X)P(Y = X \mid W = \theta) \\ &=_{\text{a.s.}} P(X - Y > c \mid Y = \theta, X > Y)P(X > Y) \\ &\quad + P(Y - X > c \mid X = \theta, Y > X)P(Y > X) \\ &=_{\text{a.s.}} P(X > \theta + c \mid Y = \theta, X > \theta)P(X > Y) \\ &\quad + P(Y > \theta + c \mid X = \theta, Y > \theta)P(Y > X) \\ &=_{\text{a.s.}} P(X > \theta + c \mid X > \theta)P(X > Y) \\ &\quad + P(Y > \theta + c \mid Y > \theta)P(Y > X) \\ h(c) &=_{\text{a.s.}} P(X > Y)[P(X > \theta + c)/P(X > \theta)] \\ &\quad + P(Y > X)[P(Y > \theta + c)/P(Y > \theta)]. \end{aligned}$$

▮

Hereafter we will abbreviate the identity of Lemma 2:

$$\text{(i)} \qquad p[f(\theta + c)/f(\theta)] + q[g(\theta + c)/g(\theta)] = h(c),$$

letting $f(a) = P(X > a)$, $g(a) = P(Y > a)$, and denote by Θ the collection of θ points whereon (i) holds for all $c \geqq 0$. Note then that if θ_n is a decreasing sequence in Θ converging to θ, then θ is in Θ, since $f(\cdot)$ and $g(\cdot)$ are right continuous. Therefore Θ contains all of its left hand end points.

Theorem 1. *Let* X *and* Y *satisfy condition* C; *if, in the above notation, there exists* $\theta_0 \varepsilon \Theta$ *such that* $(\theta_0, \theta_0 + \epsilon) \cap \Theta = \varnothing$ *some* $\epsilon > 0$, *then* X *and* Y *are discrete.*

Proof. Since $P_W(\Theta) = 1$, it follows that $P_W(\theta: \theta_0 < \theta < \theta_0 + \epsilon) = 0$, and by the corollary, $P_X(\theta: \theta_0 < \theta < \theta_0 + \epsilon) = 0 = P_Y(\theta: \theta_0 < \theta < \theta_0 + \epsilon)$.

Using (i):

$$h(c) = p[f(\theta_0 + c)/f(\theta_0)] + q[g(\theta_0 + c)/g(\theta_0)],$$

hence $h(c)$ is constant over the interval $0 < c < \epsilon$. It follows from (i) that for any $\theta \varepsilon \Theta$, $f(\cdot)$ and $g(\cdot)$ are constant over the interval $(\theta, \theta + \epsilon)$.

Thus there can be atmost countably many points of decrease for $f(\cdot)$ and $g(\cdot)$. ∎

THEOREM 2. *Let X and Y satisfy condition* C; *if the hypotheses of Theorem* 1 *are not satisfied, that is, if for every $\theta \varepsilon \Theta$, $\epsilon > 0$, $(\theta, \theta + \epsilon) \cap \Theta \neq \varnothing$, then for some constant a, $(X - a)$ and $(Y - a)$ have (possibly different) exponential distribution functions.*

PROOF. It follows from Theorem 1 and the preceding remarks that Θ is of the form $[a, +\infty)$ or $(-\infty, +\infty)$. In the former case we may adjust the right hand end point by adding a constant to X and Y; hence we may assume that $[0, +\infty) \subset \Theta$.

Now, f and g are monotone functions; hence they have right derivatives $f^{+}(\cdot)$ and $g^{+}(\cdot)$ almost everywhere. For some fixed θ, $f^{+}(\theta)$ and $g^{+}(\theta)$ are both finite, hence

$$\text{(ii)} \qquad p[f^{+}(\theta)/f(\theta)] + q[g^{+}(\theta)/g(\theta)] = h^{+}(0), \text{ finite.}$$

Since (ii) holds for all non-negative θ, it follows that $pf^{+}(\theta)$ and $qg^{+}(\theta)$ are finite for all non-negative θ, therefore the singular and discrete parts of the corresponding cumulative distribution functions must vanish; hence $pf(\cdot)$ and $qg(\cdot)$ are equal to the integral of these derivatives.

Integrating with respect to θ, $0 \leqq \theta$,

$$q \ln g(\theta) = -p \ln f(\theta) - k_2\theta + k_3, \qquad k_2 \geqq 0.$$

Hence

$$g(\theta) = f(\theta)^{-k_1} \cdot \exp[-k_2\theta + k_3], \qquad 0 \leqq \theta < r,\ k_1 \geqq 0,\ k_2 \geqq 0.$$

Going back to equation (i) we now have:

$$\text{(iii)} \qquad \begin{aligned} h(c) &= p[f(\theta + c)/f(\theta)] + q \exp(-k_2c)(f(\theta + c)/f(\theta))^{-k_1} \\ &= p[f(c)/f(0)] + q \exp(-k_2c)(f(c)/f(0))^{-k_1}. \end{aligned}$$

The equation

$$pX + q \exp(-k_2c)X^{-k_1} - p[f(c)/f(0)] - q \exp(-k_2c)(f(c)/f(0))^{-k_1} = 0$$

in X has at most two roots, (since the derivative of the curve changes sign at most once).

Therefore, if the identity (iii) in the continuous function $(f(\theta + c)/f(\theta)$ holds for all θ, then the ratio $f(\theta + c)/f(\theta)$ must be identically equal to the root $x = f(c)/f(0)$.

Hence, we have a form of Cauchy's equation:

$$f(c + \theta) = f(\theta)l(c), \qquad \theta \geqq 0,\ c \geqq 0,$$

and therefore $f(a) = k_2 \exp(-k_3a)$, $k_2, k_3 \geqq 0$, for $0 \leqq a$.

We have proved that the right hand tails of the distribution functions of X

and Y are exponential; we must show that they are exponential over their entire range. Suppose otherwise, that over the interval $(-a, -b)$, the function $f(\cdot)$ is not exponential, and $f(-a) < 1$. Then since $f(-a) < 1$,

$$P(W \leqq -a) > 0.$$

But this implies that the closed support of P_W contains a neighborhood of a point to the left of $-a$, hence f can be shown to be exponential on $[-a, +\infty)$.

Similarly it is clearly that $f(-a) < 1$ implies $g(-a) < 1$; hence X and Y have the same range, proving Theorem 2. ∎

THEOREM 3. *If X and Y are discrete random variables satisfying condition* C, *then $b(X - a)$ and $b(Y - a)$ have (possibly different) geometric distributions for suitable constants a and b.*

PROOF. It follows from Theorems 1 and 2 that there exists θ_0 such that $\theta_0 + \epsilon \not\in \Theta$ for all small ϵ, hence there exists a smallest c_0 such that $\theta_0 + c_0$ is a point of decrease for $f(\cdot)$ or $g(\cdot)$. Now

$$p[f(\theta_0 + c)/f(\theta_0)] + q[g(\theta_0 + c)/g(\theta_0)] = h(c).$$

Thus c_0 is the smallest positive point of decrease for $h(\cdot)$. Hence, it follows from (i) that for any $\theta \in \Theta$, $\theta' = \theta + c_0$ is the next point of decrease for $f(\cdot)$ or $g(\cdot)$, and $\theta'' = \theta' + c_0$ is the next, and so on. Thus the support of P_W is a right-unbounded collection of lattice points. If necessary, we make an affine transformation and assume they are a subset of the integers containing the non-negative integers.

Using condition C we may write

$$P(X - Y > m)P(W = n) = P(X - Y > m, W = n)$$
$$= P(X - Y > m, X > Y = n) = P(X > m + n)P(Y = n).$$

Let $r(m) = \lg P(X > m)$. Then $r(\cdot)$ satisfies an equation of the form

$$r(m + n) = s(m) + t(n)$$
$$r(m + 1) - r(m) = t(1) - t(0)$$
$$r(m) = r(0) + m(t(1) - t(0)).$$

Thus $P(X > m) = \delta p^m$ for some δ, p. It follows that $0 < \delta; 0 < p < 1$; that is, the right hand tail of the distribution of X is geometric. Similarly, the right hand tail of the distribution of Y is geometric; and the same argument used in Theorem 2 suffices to complete Theorem 3. ∎

3. Acknowledgment. The author is indebted to the referee for the reference [4], as well as many helpful comments.

REFERENCES

[1] BASU, A. P., (1965). On characterizing the exponential distribution by order statistics. *Ann. Inst. Statist. Math.* **17** 93–96.

[2] FERGUSON, T. S., (1965). A characterization of the geometric distribution. *Amer. Math. Monthly* **72** 256–260.

[3] FERGUSON, T. S., (1964). A characterization of the exponential distribution. *Ann. Math. Statist.* **35** 1199–1207.

[4] ROGERS, G. S., (1963). On alternate proof of the characterization of the density Ax^B. *Amer. Math. Monthly* **70** 857–858.