The Annals of Mathematical Statistics
1970, Vol. 41, No. 6, 2113–2122

A CHARACTERIZATION BASED ON THE ABSOLUTE DIFFERENCE OF TWO I.I.D. RANDOM VARIABLES

BY PREM S. PURI[1] AND HERMAN RUBIN[2]

Purdue University

1. Introduction. Let X_1 and X_2 be two independent and identically distributed (i.i.d.) random variables whose common distribution is the same as that of a random variable X. The problem considered here is to characterize all possible distributions of X which satisfy the following property H:

$$H: \text{ The distribution of } |X_1 - X_2| \text{ and } X \text{ are identical.} \tag{1}$$

For instance, it is easy to verify that the discrete distribution with $P(X=0) = P(X=a) = \frac{1}{2}$ for some positive constant a, and the exponential distribution with probability density function (pdf) f where $f(x) = \theta \exp(-\theta x)$, for $x \geqq 0$, and $f(x) = 0$ elsewhere, with $\theta > 0$, both satisfy the property H. The reader may find a different characterization based on $|X_1 - X_2|$ in Puri [6]. Basu [1], Ferguson ([4], [5]) and Crawford [2] have considered a different problem where they characterize distributions with the property that $\min(X_1, X_2)$ is independent of $X_1 - X_2$. Their methods naturally depend very heavily upon such an independence, which of course is lacking in the present case.

Let F denote the distribution function (df) of X. It can be easily shown that if X satisfies H, the distribution of X can either be only discrete or absolutely continuous or singular and no mixture is possible. Thus one needs to consider these three possibilities separately. For the case when X is discrete let A denote the set of possible discrete nonnegative values that X takes. More specifically, let

$$p_y = P(X=y), \qquad y \in A; \quad \text{with} \quad \textstyle\sum_{y \in A} p_y = 1.$$

It is clear that if there exists a $y \geqq 0$ with $p_y > 0$, then in particular A contains zero with $p_0 > 0$. Furthermore, from the property H, the following relations follow easily:

$$p_0 = \textstyle\sum_{x \geqq 0} p_x^{\,2}, \tag{2}$$

$$p_y = 2 \textstyle\sum_{x \geqq 0} p_x p_{x+y}; \qquad y > 0. \tag{3}$$

Similar relations are satisfied by the pdf f if X satisfying H is absolutely continuous.

In Section 2, we show that under H, X has a moment generating function (mgf) and hence all its moments are finite. Also in Theorem 1, we consider the case

Received January 26, 1970; revised June 3, 1970.

[1] This investigation was supported in part by research grant GM-10525 from NIH, Public Health Service, at the University of California, Berkeley.

[2] This research was partly supported by the Office of Naval Research Contract N00014-67-A-0226-0008, project NR042-216. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

where X is bounded. Section 3 deals with the discrete case, and Theorem 2 characterizes lattice distributions satisfying H. In Section 4, we consider the absolutely continuous case. Here we study a more general question; namely, if X_1 and X_2 are two nonnegative independent but not necessarily identically distributed random variables (rv), and moreover if the distributions of X_1 and $|X_1 - X_2|$ are identical, then given the distribution of X_2, what can be said about the distribution of X_1? The paper ends with a discussion in Section 5, where we have a few words to say about the singular case.

2. Preliminary results. In the following lemma it is shown that for an X satisfying H, its mgf and hence all its moments exist.

LEMMA 1. *The* mgf *of a* rv *X satisfying H exists.*

PROOF. If X satisfying H is degenerate, it is clear that $P(X=0)=1$, and the lemma holds trivially. Let X be nondegenerate. Then there is a number $u > 0$ such that $P(X \leqq u) > \frac{1}{2}$. Using this and the property H, it follows that for every $v \geqq 0$

$$P(X > v) = P(|X_1 - X_2| > v) \geqq 2P(X > u+v)P(X \leqq u), \tag{4}$$

so that

$$P(X > u+v) \leqq P(X > v)/2P(X \leqq u),$$

for all $v \geqq 0$. A repeated application of this leads to

$$P(X > nu) \leqq \left[\frac{1}{2P(X \leqq u)}\right]^{n-1}; \qquad \text{for} \quad n = 1, 2, \cdots.$$

From this one can easily show the existence of an $\alpha > 0$ such that $E(\exp(\xi X))$ exists for all $|\xi| \leqq \alpha$.

The following theorem provides the answer to our problem when X is bounded.

THEOREM 1. *Let X be nondegenerate. Then the following three statements are equivalent.*

(i) *X is bounded and satisfies H.*
(ii) *X satisfies H and $P(X=0) = \frac{1}{2}$.*
(iii) *$P(X=0) = P(X=a) = \frac{1}{2}$, for some $a > 0$.*

PROOF. Clearly (iii) $\Rightarrow$ (i) and (ii). All we need to prove is that (i) $\Rightarrow$ (iii) and (ii) $\Rightarrow$ (iii). Let (i) hold. Since X is bounded and nondegenerate, there exists a least upper bound $B > 0$, such that $P(X > B) = 0$ and for every $0 < \varepsilon \leqq B$, $P(X > B - \varepsilon) > 0$. On the other hand, since X satisfies H, we have for every such ε

$$0 < P(X > B-\varepsilon) = P(|X_1 - X_2| > B-\varepsilon) \leqq 2P(X < \varepsilon)P(X > B-\varepsilon), \tag{5}$$

which implies that for every $0 < \varepsilon \leqq B$, $P(X < \varepsilon) \geqq \frac{1}{2}$ and, in particular, letting ε tend to zero we have $P(X=0) \geqq \frac{1}{2}$. This implies that X must be a discrete random variable (rv). Using the notation introduced in Section 1 for such a case, we have $\frac{1}{2} \leqq p_0 = \sum_{x \geqq 0} p_x^2 < 1$. This yields

$$p_0(1-p_0) = \sum_{x>0} p_x^2 \leqq (\max_{x>0} p_x)(\sum_{x>0} p_x) = (\max_{x>0} p_x)(1-p_0),$$

so that $p_0 \leqq \max_{x>0} p_x$. Hence we have

$$\frac{1}{2} \leqq p_0 \leqq \max_{x>0} p_x \leqq 1 - p_0 \leqq \frac{1}{2}, \tag{6}$$

so that $p_0 = \max_{x>0} p_x = \frac{1}{2}$, which implies that $P(X=0) = P(X=B) = \frac{1}{2}$. This proves that (i) ⇒ (iii) with $B = a$. Now let (ii) hold. Since $p_0 = \frac{1}{2}$, X must be a discrete rv if it has to satisfy H. For this we have already seen that $p_0(1-p_0) = \sum_{x>0} p_x^2$. This means that we have

$$\sum_{x>0} p_x^2 = \frac{1}{4} \quad \text{and} \quad \sum_{x>0} p_x = \frac{1}{2}. \tag{7}$$

But this holds if and only if $p_x = \frac{1}{2}$ for some $x = a > 0$, so that (iii) holds. □

Before closing this section we wish to remark that for the nondegenerate discrete case, for X satisfying H we must have $0 < p_0 \leqq \frac{1}{2}$; that $p_0 > 0$ follows from (2), and the fact that $\sum_{x \geqq 0} p_x = 1$; that $p_0 \leqq \frac{1}{2}$ follows from the fact that for every $y > 0$ with $p_y > 0$, $p_y \geqq 2p_0 p_y$ under H.

3. Discrete case. We now consider the case where X is discrete and satisfies the following additional condition C.

$$\text{C: There exists an interval } (\delta_1, \delta_2] \text{ with } 0 \leqq \delta_1 < \delta_2 < \infty \text{ such that } P(\delta_1 < X \leqq \delta_2) = 0. \tag{8}$$

Using the notation of Section 1, we first prove three lemmas needed to prove the main result of Theorem 2.

LEMMA 2. *Let X be discrete, nondegenerate and satisfy H and the condition* C. *Then*

(i) $\tau = \inf\{x : x > 0, p_x > 0\} > 0$ *and* $p_\tau > 0$, *and*
(ii) *the set of possible values of X is given by* $k\tau$; $k = 0, 1, 2, \cdots$.

PROOF. (i) Since X is nondegenerate, it is clear that the set $\{x : x > 0, p_x > 0\}$ is nonempty. Again, if $p_0 = \frac{1}{2}$, (i) and (ii) are satisfied in view of Theorem 1, so that let $0 < p_0 < \frac{1}{2}$. By Theorem 1, this means that X is not bounded. Let S be the set of possible values of X. In view of the property H, it is easy to show that S forms a positive linear space in integers. By this we mean that if $x_i \in S$, $i = 1, 2, \cdots$, then $|\sum_i n_i x_i| \in S$ for all integer values (positive or negative) of n_i's. Now for an unbounded set with this property, it is not difficult to show that either this set is dense everywhere over $[0, \infty)$ or is a lattice. On the other hand, in view of condition C, it cannot be dense everywhere. Hence the lemma follows. □

In view of Lemma 2, let p_k denote the probability $\Pr(X = k\tau)$, for $k = 0, 1, 2, \cdots$, so that $\sum_{k=0}^{\infty} p_k = 1$. The analogues of (2) and (3) are given by

$$p_0 = \sum_{i=0}^{\infty} p_i^2, \tag{9}$$

$$p_k = 2\sum_{i=0}^{\infty} p_i p_{i+k}; \qquad k = 1, 2, \cdots. \tag{10}$$

We shall now restrict to the case with $0 < p_0 < \frac{1}{2}$. Thus the set of possible values of X, in view of Theorem 1, must be infinitely denumerable. Furthermore, from (12) of the following lemma it follows that under condition C, $p_k > 0$, for $k = 1, 2, \cdots$.

LEMMA 3. *Let H and* C *hold and also let* $0 < p_0 < \frac{1}{2}$. *Then for* $k = 1, 2, \cdots$,

$$p_k \geqq \left(\frac{2p_1}{1-2p_0}\right) \cdot p_{k+1} \tag{11}$$

and

$$p_{k+1} \geqq \frac{2p_1(1-2p_0)}{[(1-2p_0)^2+4p_1{}^2]} \cdot p_k. \tag{12}$$

PROOF. (11) follows easily from (10) by noticing that for $k \geqq 1$,

$$p_k(1-2p_0)-2p_1\, p_{k+1} = 2\textstyle\sum_{i=2}^{\infty} p_i\, p_{i+k}, \tag{13}$$

and that the right side of (13) is nonnegative. To prove (12), we first notice from (10) that for $k \geqq 1$,

$$p_{k+1} = 2\textstyle\sum_{i=0}^{\infty} p_i\, p_{i+k+1}, \tag{14}$$

or equivalently

$$p_{k+1}(1-2p_0) = 2\textstyle\sum_{i=1}^{\infty} p_i\, p_{i+k+1}. \tag{15}$$

Then using (11) for each p_i on the right side of (15) we have

$$\begin{aligned} p_{k+1}(1-2p_0) &\geqq \frac{4p_1}{1-2p_0}\sum_{i=1}^{\infty} p_{i+1}\, p_{i+k+1} = \frac{2p_1}{1-2p_0}\left(2\sum_{i=2}^{\infty} p_i\, p_{i+k}\right) \\ &= \frac{2p_1}{1-2p_0}(p_k - 2p_0\, p_k - 2p_1\, p_{k+1}). \end{aligned} \tag{16}$$

Here at the end of (16) we have again used (10). Finally (12) follows immediately from (16) after a little simplification. □

For each sequence $\{p_k\}$ satisfying H (or equivalently (9) and (10)) and with $0 < p_0 < \frac{1}{2}$, define

$\beta = \sup\{b : b > 0$ satisfying $p_k \geqq bp_{k+1}$ for all $k \geqq 1\}$ and

$\gamma = \sup\{c : 0 < c < 1$, satisfying $p_{k+1} \geqq cp_k$ for all $k \geqq 1\}$, so that

$$p_k \geqq \beta p_{k+1}; \qquad k = 1, 2, \cdots, \tag{17}$$

and

$$p_{k+1} \geqq \cdot \gamma p_k; \qquad k = 1, 2, \cdots. \tag{18}$$

From Lemma 3 it follows that for every sequence $\{p_k\}$ satisfying H and with $0 < p_0 < \frac{1}{2}$, there always exist positive β and γ. Also by definition of β, it is clear that for every such sequence $0 < \beta \leqq p_1/p_2$. Here $p_2 > 0$; in fact because of (12)

$p_k > 0$ for all $k \geqq 1$. Also γ has to be strictly between 0 and 1. That it cannot be equal to one follows from (18) and the fact that $\sum_{i=0}^{\infty} p_i$ converges. The following lemma is the essential lead to the main theorem of this section.

LEMMA 4. *For every sequence* $\{p_k\}$ *with* $0 < p_0 < \frac{1}{2}$ *and satisfying* $H, \beta\gamma = 1$.

PROOF. From (17) and (18) it is clear that $\beta\gamma \leqq 1$. It is sufficient then to prove that $\beta\gamma \geqq 1$. From (14) we have for $k \geqq 2$,

$$p_k(1-2p_0) = 2\textstyle\sum_{i=1}^{\infty} p_i p_{i+k}. \tag{19}$$

Using (18) on the right side of (19), for each p_i we have

$$\begin{aligned} p_k(1-2p_0) &\leqq \frac{2}{\gamma}\sum_{i=1}^{\infty} p_{i+1} p_{i+k} = \frac{2}{\gamma}\sum_{i=2}^{\infty} p_i p_{i+k-1} \\ &= \frac{1}{\gamma}\left[p_{k-1} - 2p_0 p_{k-1} - 2p_1 p_k\right], \end{aligned} \tag{20}$$

which after simplification, yields for $k = 2, 3, \cdots$,

$$p_{k-1} \geqq \left[\gamma + \frac{2p_1}{1-2p_0}\right] p_k,$$

or equivalently for $k = 1, 2, \cdots$,

$$p_k \geqq \left[\gamma + \frac{2p_1}{1-2p_0}\right] p_{k+1}. \tag{21}$$

Comparing (17) and (21) and keeping the definition of β in mind, we have

$$(\beta - \gamma) \geqq \frac{2p_1}{1-2p_0}. \tag{22}$$

Again, using (17) on the right side of (19), we have for $k \geqq 2$,

$$\begin{aligned} p_k(1-2p_0) &\geqq 2\beta\textstyle\sum_{i=1}^{\infty} p_{i+1} p_{i+k} = \beta(2\textstyle\sum_{i=2}^{\infty} p_i p_{i+k-1}) \\ &= \beta[p_{k-1} - 2p_0 p_{k-1} - 2p_1 p_k]. \end{aligned} \tag{23}$$

On simplification, (23) yields for $k \geqq 2$,

$$p_k \geqq \frac{\beta(1-2p_0)}{(1-2p_0+2p_1\beta)} p_{k-1},$$

or equivalently for $k \geqq 1$,

$$p_{k+1} \geqq \frac{\beta(1-2p_0)}{(1-2p_0+2p_1\beta)} p_k. \tag{24}$$

Finally comparing (18) and (24) and using the definition of γ, we obtain

$$\gamma \geqq \frac{\beta(1-2p_0)}{(1-2p_0+2p_1\beta)}$$

or after simplification

$$(\beta-\gamma) \leqq \frac{2p_1}{1-2p_0}\beta\gamma. \tag{25}$$

Now it easily follows from (22) and (25) that $\beta\gamma \geqq 1$. □

We are now in a position to state and prove the main theorem of this section.

THEOREM 2. *Let X_1 and X_2 be two independent copies of a nonnegative discrete random variable X satisfying condition* C. *Then X and the absolute difference $|X_1-X_2|$ have the same distribution if and only if the distribution of X is given for some positive constant τ, by*

$$\begin{aligned} &\Pr(X = 0) = p_0 \\ &\Pr(X = k\tau) = 2p_0(1-p_0)(1-2p_0)^{k-1}; \qquad k = 1,2,\cdots, \end{aligned} \tag{26}$$

where either $p_0 = 1$ or $0 < p_0 \leqq \frac{1}{2}$.

PROOF. The case with $p_0 = 1$ is that of a degenerate rv X. Also we have argued before that for a nondegenerate X, we must have $0 < p_0 \leqq \frac{1}{2}$. The case with $p_0 = \frac{1}{2}$ is covered in Theorem 1. Let us assume then that $0 < p_0 < \frac{1}{2}$. From Lemma 4 and equations (17) and (18) it follows that

$$p_{k+1} = \gamma p_k; \qquad k = 1,2,\cdots, \tag{27}$$

or equivalently

$$p_k = \gamma^{k-1}p_1; \qquad k = 1,2,\cdots. \tag{28}$$

Now it is easy to show using (9) and the fact that $\sum_{i=0}^{\infty} p_i = 1$, that $\gamma = (1-2p_0)$ and $p_1 = 2p_0(1-p_0)$. □

4. Absolutely continuous case. Let $f(x)$ denote the pdf of the nonnegative rv X with the property H. The property H is then equivalent to $f(x)$ satisfying the relations

$$\int_0^\infty f(x)\,dx = 1; \qquad f(t) = 2\int_0^\infty f(x+t)f(x)\,dx; \qquad \text{for all} \quad t \geqq 0. \tag{29}$$

Furthermore, in view of Theorem 1, X is unbounded. The following lemma gives certain properties of an f satisfying (29), which we shall need later.

LEMMA 5. *Let the* pdf $f(x)$ *of a nonnegative* rv X *satisfy* (29). *Then it also satisfies the following:*

(i) *$f(x)$ is lower semicontinuous for all $x \geqq 0$.*
(ii) *$f(x) > 0$, for all $x \geqq 0$.*

PROOF. (i) From (29) for $t = 0$, we have

$$f(0) = 2\int_0^\infty f^2(x)\,dx; \qquad \int_0^\infty f(x)\,dx = 1, \tag{30}$$

so that we must have $f(0) > 0$. On the other hand, since

$$\psi(z) = \int_0^\infty f(y+z)f(y)\,dy; \qquad -\infty < z < \infty,$$

is the pdf of $X_1 - X_2$, or equivalently the pdf of the convolution of X_1 and $-X_2$, $\psi(z)$ is lower semicontinuous for all $-\infty < z < \infty$. The last statement follows from Fatou's Lemma and the measurability of f. By virtue of (29) therefore, $f(t)$ is lower semicontinuous for all $t \geqq 0$.

(ii) Assume that there exists an interval (a, b) with $\varepsilon < a < b$, such that $f(x) = 0$ for all $x \in (a, b)$. Since $f(0) > 0$ and $f(x)$ is lower semicontinuous at zero, there is an $\varepsilon > 0$, with $a < b - \varepsilon/2$, such that $f(x) > 0$, for all $x \in [0, \varepsilon]$. Using this and (29), it is now easy to show that $f(z) = 0$, for all $a < z < b + \varepsilon/2$, a.e. μ. By an induction argument we then have $f(z) = 0$ for all $a < z < b + n\varepsilon/2$, a.e. μ, for $n = 1, 2, \cdots$. Letting $n \to \infty$, we have $f(z) = 0$ for all $a < z < \infty$, a.e. μ. But this implies that X is bounded, which is a contradiction. Thus there exists no interval $I \subset [0, \infty)$ with $\mu(I) > 0$ such that $f(x) = 0$ for all $x \in I$. This implies that $f(x) > 0$ for all $x \geqq 0$, a.e. μ. Now let $f(x_0) = 0$ for some $x_0 > 0$. Then

$$\int_0^\infty f(x+x_0) f(x)\, dx = 0 \Rightarrow \int_a^b f(x+x_0) f(x)\, dx = 0, \quad \text{for } 0 < a < b < \infty,$$
$$\Rightarrow f(x+x_0) = 0, \quad \text{for all } a < x < b \text{ with } f(x) > 0, \text{ a.e. } \mu.$$

But $\mu[x : a < x < b, f(x) > 0] = b - a$, which also yields $\mu[x + x_0 : a < x < b, f(x) > 0] = b - a > 0$. This contradicts the fact that $f(x) > 0$, for all $x \geqq 0$ a.e. μ. Thus $f(x) > 0$, for all $x \geqq 0$. □

We shall now consider a more general problem. Let X and Y be two nonnegative independently but not necessarily identically distributed random variables. Let F and G denote the df's of X and Y respectively. Given F and that the distributions of Y and $|Y - X|$ are identical, what can we say about the distribution of Y, i.e. about G? The reader may find in Feller ([3], pages 208–209) a treatment of this problem considered for a somewhat restricted case. That the distributions of Y and $|Y - X|$ are identical is equivalent to the relation

$$G(t) = \int_0^\infty G(x+t)\, dF(x) + \int_0^\infty F(y+t)\, dG(t); \qquad \text{for all} \quad t \geqq 0. \tag{31}$$

The following theorem provides an answer to the question raised above.

THEOREM 3. *Let X and Y be two nonnegative independent random variables with F and G as their respective* df*'s. Let $EX < \infty$ and F have an absolutely continuous part. Then G satisfies* (31) *if and only if G is absolutely continuous with* pdf *g where*

$$g(y) = [1 - F(y)]/EX; \qquad \textit{for all} \quad y \geqq 0. \tag{32}$$

PROOF. It is easy to verify that $g(y)$ of (32) does satisfy (31). All we need to show is that this is the unique g that satisfies (31). To this end, consider a sequence of i.i.d. random variables $X_1, X_2, X_3, \cdots$, with their common distribution same as that of X. Define another sequence of random variables Z_k recursively by

$$Z_1 = X_1, \qquad Z_{n+1} = |Z_n - X_{n+1}|. \tag{33}$$

Then clearly $\{Z_n\}$ is a Markov chain (MC) with state space $[0, \infty)$ and the transition df H, given by

$$dH(x \mid y) = dF(x-y) + dF(x+y), \tag{34}$$

so that if G_n is the df of Z_n, it is easily observed that $G_1 = F$ and for $n = 2, 3, \cdots$,

$$G_n(y) = \int_0^\infty G_{n-1}(x+y)\,dF(x) + \int_0^\infty F(x+y)\,dG_{n-1}(x). \tag{35}$$

Letting $n \to \infty$, we observe that any solution G of (31) is a stationary distribution of the MC $\{Z_n\}$. We have already observed that $g(y)$ given by (32) is such a stationary distribution. That this is the unique stationary distribution and hence the unique solution of (31) follows from the fact that the above MC defined on $[0, \infty)$ is indecomposable, which is a simple consequence of the fact that F has an absolutely continuous part.

In answer to our original question, we now have the following theorem.

THEOREM 4. *Let X_1 and X_2 be two independent copies of a* rv *X with* pdf $f(x)$. *Then X and $|X_1 - X_2|$ have the same distribution, if and only if for some* $\theta > 0$,

$$\begin{aligned} f(x) &= \theta e^{-\theta x}, \quad && \textit{for} \quad x \geqq 0 \\ &= 0 && \textit{elsewhere}. \end{aligned} \tag{36}$$

PROOF. Clearly if $f(x)$ satisfies (36), the distributions of X and $|X_1 - X_2|$ are identical. Assuming now that the distributions of X and $|X_1 - X_2|$ are the same, it is easily seen that $f(x)$ satisfies the conditions of Theorem 3, in view of Lemma 1 and Lemma 5. On the other hand, comparing (29) and (31), we have under H, $g(y) = f(y)$, so that replacing $g(y)$ with $f(y)$ in (32), and solving the resulting equation for $f(y) = F'(y)$ we obtain (36) with $\theta = EX$. □

5. A few concluding remarks. The lines of proof adopted for Theorem 3, and hence of Theorem 4, in principle should also work for the discrete case of Section 3. Let X and Y be two appropriate nonnegative discrete rv, both independently but not necessarily identically distributed with $\{p_x\}$ and $\{p_y\}$ as the set of their probabilities (as defined in Section 1). Given that the distributions of Y and $|Y - X|$ are identical, the analogue of equation (29) is given by

$$\begin{aligned} q_0 &= \textstyle\sum_{x \geqq 0} p_x q_x \\ q_y &= \textstyle\sum_{x \geqq 0} p_x q_{x+y} + \sum_{x \geqq 0} q_x p_{x+y}; \qquad \text{for} \quad y > 0. \end{aligned} \tag{37}$$

However, here essentially it is a matter of first guessing a general solution of (37) for q_y's satisfying $\sum_{y \geqq 0} q_y = 1$, in terms of p_x's, an analogue of (32). After this, replacing q_y's with p_x's in this solution, p_y's can be explicitly obtained to yield the answer to our original problem.

Concerning the singular case of an X with property H, at present we can only say in view of Theorem 1, that X has to be unbounded. On the other hand, let us consider again the approach adopted in Section 4. Let $F(x)$ and $G(y)$ respectively be the continuous distribution functions of two nonnegative independent random variables X and Y. This will cover both absolutely continuous and singular cases of our problem. Introduce a MC similar to the one of Section 4, defined on $[0, \infty)$, but with the assumption that X has the continuous df $F(x)$, so that (31) and

(35) are still satisfied. Any solution of (31) is a stationary df G of our MC $\{Z_n\}$. On the other hand, if $EX < \infty$, it is easily verifiable that

$$dG(y)/dy = [1-F(x)]/EX; \qquad y \geqq 0, \tag{38}$$

is a solution of (31) and hence a stationary df of MC $\{Z_n\}$. The only problem here is to show that (38) is the unique solution of (31). For this we need to show that the MC $\{Z_n\}$ is indecomposable. Once this is established, (38) is the unique solution of (31). The solution to our problem is then obtained by replacing G with F in (38) and solving this for F. This turns out to be the same as (36). Thus, subject to the uniqueness of the solution of (31), the solution to our original problem would be (36) even when the df of X is given to be only continuous. This would mean that there is no singular distribution with the property H. Our conjecture is that this is in reality the case.

In Section 3 for the discrete case the result of Theorem 2 was proved subject to the condition C. Our conjecture is that this result holds even without this extra condition.

Again, as suggested by a referee (see also Rogers [7]), it is worth noticing that the property H can be transformed by taking $U_i = \exp[-X_i]$, $i = 1, 2$, to

(39) H': The distribution of $\min(U_1/U_2, U_2/U_1)$ and U are identical, where U_1 and U_2 are i.i.d. random variables whose common distribution is the same as that of a positive random variable U. Clearly in order that U satisfies H', we must have $0 < U \leqq 1$. The condition C now takes the form

(40) C': There exists an interval $[\Delta_1, \Delta_2)$ with $0 < \Delta_1 < \Delta_2 \leqq 1$, such that $\Pr(\Delta_1 \leqq U < \Delta_2) = 0$.

As before, a rv U satisfying H' can either be only discrete or absolutely continuous or singular and no mixture is possible. The following Corollaries 1 and 2 follow easily now from Theorem 2 and Theorem 4 respectively.

COROLLARY 1. *Let U_1 and U_2 be two independent copies of a positive discrete random variable U satisfying condition C′. Then U and* $\min(U_1/U_2, U_2/U_1)$ *have the same distribution if and only if the distribution of U is given for some constant $0 < a < 1$, by*

$$\begin{aligned} &\Pr(U = 1) = p_0 \\ &\Pr(U = a^k) = 2p_0(1-p_0)(1-2p_0)^{k-1}; \qquad k = 1, 2, \cdots, \end{aligned} \tag{41}$$

where either $p_0 = 1$ or $0 < p_0 \leqq \frac{1}{2}$.

COROLLARY 2. *Let U_1 and U_2 be two independent copies of a positive random variable U with* pdf $h(u)$. *Then U and* $\min(U_1/U_2, U_2/U_1)$ *have the same distribution, if and only if for some $\theta > 0$,*

$$\begin{aligned} h(u) &= \theta u^{\theta-1}, \qquad 0 < u \leqq 1 \\ &= 0 \qquad \textit{elsewhere}. \end{aligned} \tag{42}$$

Finally, if X_1 and X_2 are i.i.d. with df F, consider the property

(43) H^*: The df of $X_1 - X_2$ is given by $\frac{1}{2}[F(x)+1-F(-x-)]$.

It is easy to see that when X is nonnegative, H^* is equivalent to H of (1). Also, it is easy to show that Theorem 1 still holds when H is replaced by H^* and the constant a is allowed to be either strictly positive or negative. This would cover the case when X is bounded. For the unbounded case there appear to be several distributions with the property (43). In particular, in answer to a referee's question whether there can be an X which has a pdf and satisfies H^*, but is not an exponential rv, we give below such an example. For the case, where F is continuous, H^* is equivalent to

$$\mathscr{R}(\varphi(t)) = |\varphi(t)|^2, \tag{44}$$

where $\mathscr{R}(\varphi(t))$ is the real part of the characteristic function φ of X. Let X have the pdf given by

$$\begin{aligned} f(x) &= \frac{(1-\beta)}{(1-2\beta)(1+\beta)}e^{-x} + \frac{(1-\beta)}{(2\beta-1)(2-\beta)}\exp\left(-\left(\frac{1-\beta}{\beta}\right)x\right), & x > 0 \\ &= \frac{(1-\beta)}{(1+\beta)(2-\beta)}e^{x/\beta}, & x < 0, \end{aligned} \tag{45}$$

where $0 < \beta < 1$. The characteristic function of this is given by

$$\varphi(t) = [(1-it)(1-(\beta it/1-\beta))(1+\beta it)]^{-1}, \tag{46}$$

which can be easily shown to satisfy (44).

REFERENCES

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

[2] Crawford, G. B. (1966). Characterization of geometric and exponential distributions. *Ann. Math. Statist.* **37** 1790–1795.

[3] Feller, W. (1966). *An Introduction to Probability Theory and its Applications*, **2**. Wiley, New York.

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

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

[6] Puri, Prem S. (1966). Probability generating functions of absolute difference of two random variables. *Proc. Nat. Acad. Sci. USA* **56** 1059–1061.

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